



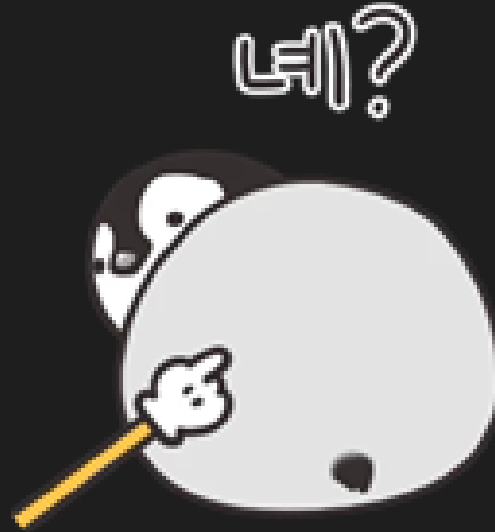
Web Crawling 기초 및 실습

Second Session

엄펑 쓰앵님



오늘은....
더 어려워요 여러분...



하지만 오늘도 역시 **질문**합니다
대답 못하시면 ... ㅎㅎ..
긴장 바짝하고 집중 하셔야해용

수강생 소개



수원



진Aㅏ



어쩔다지5

가이니

부탁이 있어요



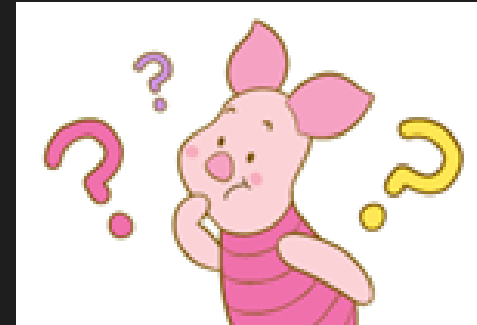
문짚왕 은진띠



쿼터 구



비된찌



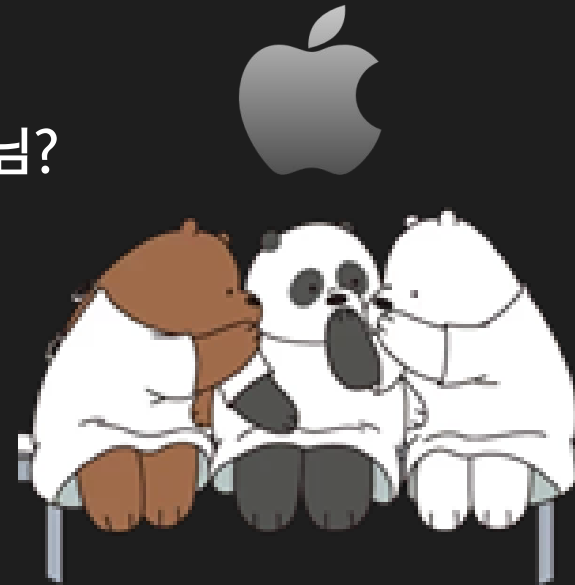
또글렛

그런데



엄펍 쌤님

누가 요새 맥북 아님?
쑥덕쑥덕



쑥덕쑥덕

수강생 여러분

오늘도 맥에서는 오류가 나지 않을까...
매우 떨립니다..

뿌이뿌이뿌이~



그럼 시작해 볼까요
준비 되셨나요오오

오늘은 어떤 것을 배울까요옹

네이버 영화 평점 / 리뷰

네이버 지도 즐겨찾기

네이버 웹툰 베댓

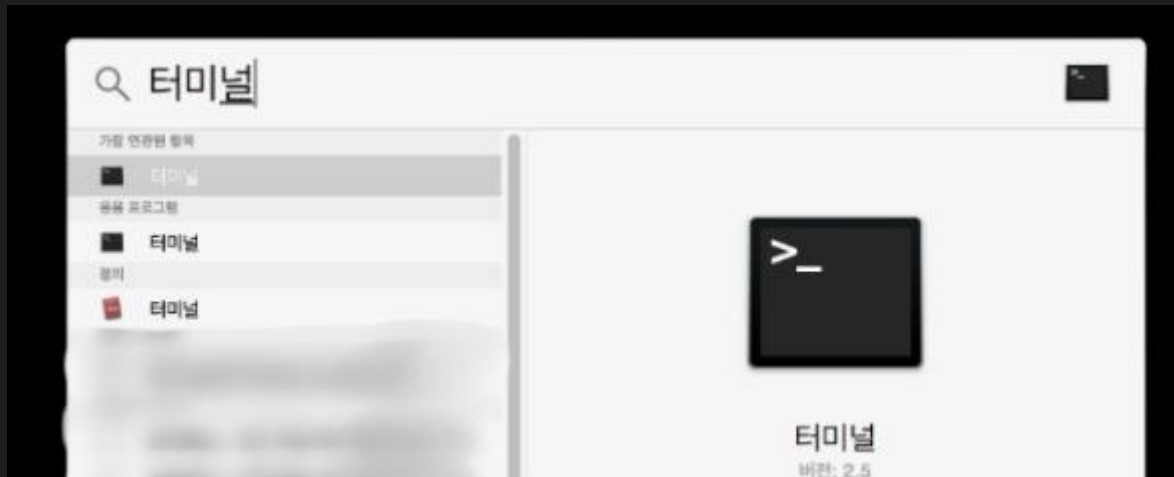


환경 셋팅, 설치

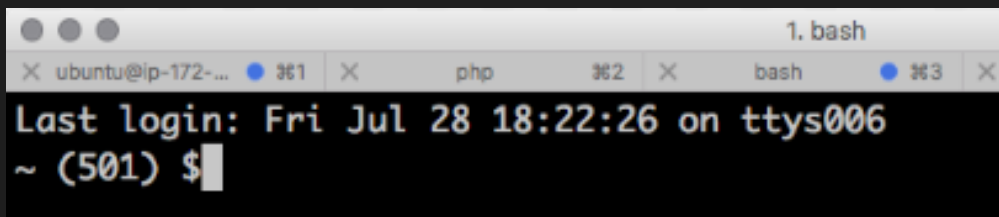


Spotlight 검색

cmd + 스페이스로 spotlight 켜기



‘터미널’ 혹은 ‘Terminal’ 검색



웹을 제어하는 **selenium**이라는 라이브러리 다운

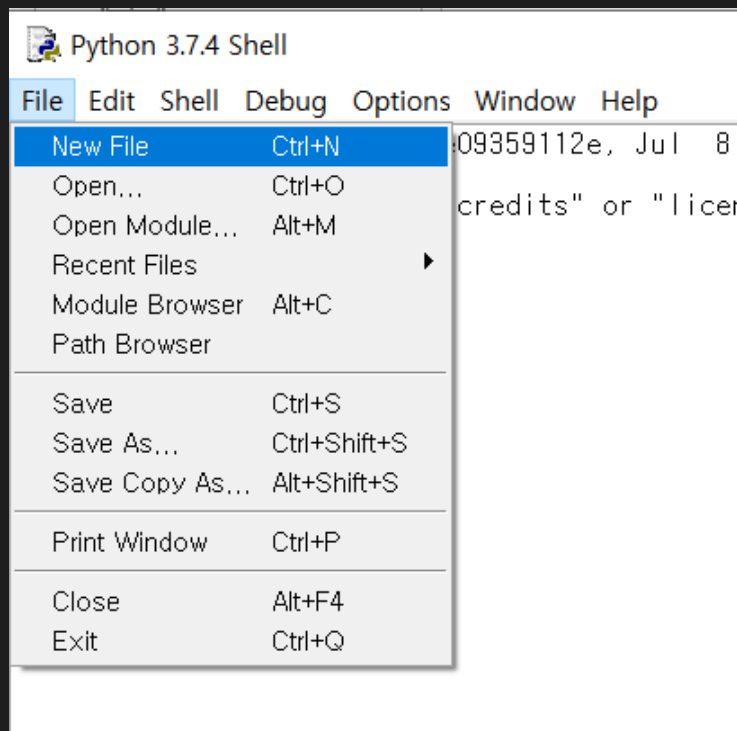
```
$ python3 -m pip install --user selenium
```


환경 셋팅, 설치

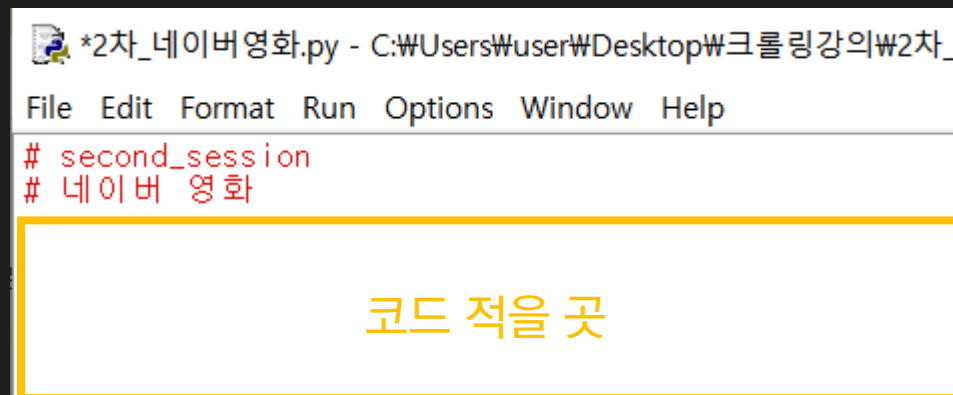
Python IDLE 실행



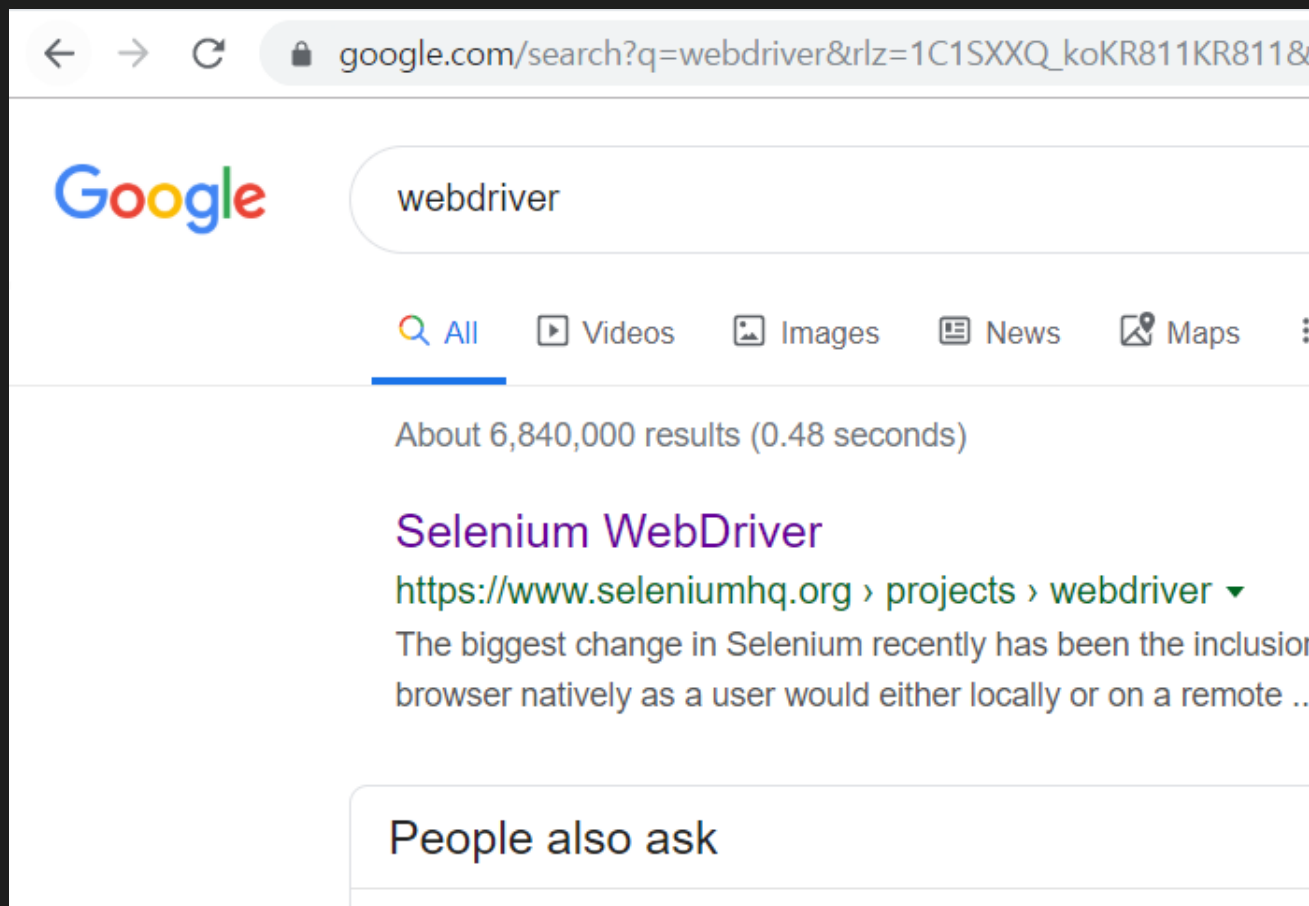
File → New File



다른 이름으로 저장 후 주석 달기



환경 셋팅, 설치



Chrome webdriver 다운

Apache License v.2.0. Some of the plugins are available under the Apache License; others are only available under a proprietary license. The license of distribution need to be raised with their respective owners.

Third Party Browser Drivers **NOT DEVELOPED** by seleniumhq

Browser

Mozilla GeckoDriver	latest	change log
Google Chrome Driver	latest	change log
Opera	latest	
Microsoft Edge Driver		
GhostDriver	(PhantomJS)	

<https://www.seleniumhq.org/download/>

ChromeDriver






WebDriver is an open source tool for automated testing of webapps by simulating a user navigating to web pages, user input, JavaScript execution, and more. It implements the [W3C WebDriver standard](#). ChromeDriver is available for Chrome on Windows, Linux, and ChromeOS).

You can view the current implementation status of the WebDriver

All versions available in Downloads

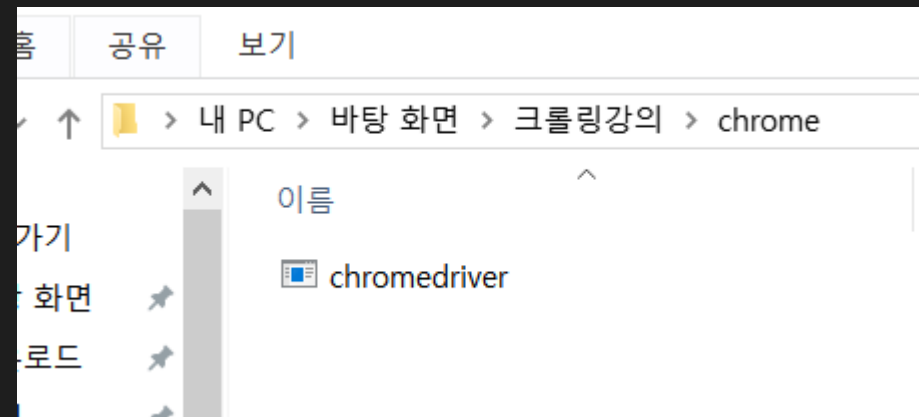
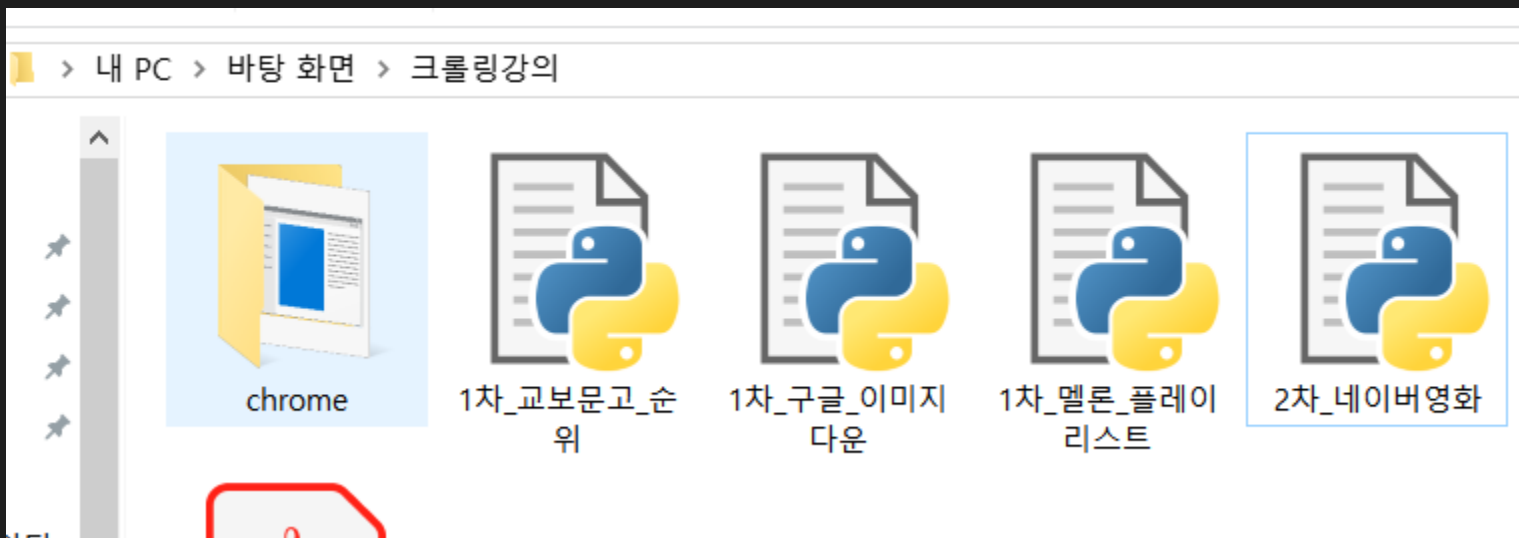
- Latest stable release: [ChromeDriver 77.0.3865.40](#)
- Latest beta release: [ChromeDriver 78.0.3904.11](#)

Index of /77.0.3865

	Name	Last Modified
	Parent Directory	
	chromedriver_linux64.zip	2019-0
	chromedriver_mac64.zip	2019-0
	chromedriver_win32.zip	2019-0
	notes.txt	2019-0

Mac용 크롬 드라이버 다운

환경 셋팅, 설치



네이버지도 영화 파일이 있는 곳에 압축풀기

Q. Selenium? Webdriver?

The screenshot displays a development environment with the following components:

- Code Editor (Brackets):** Shows the file explorer on the left with a project structure including `section02-1-c.py` through `section04-1-c.py`. The main editor displays the code for `section06-2-c.py`, which is a Selenium WebDriver script for navigating to Naver and performing actions.
- Terminal Window:** A terminal window titled "2차_네이버지도_즐거찾기.py - C:\Users\User\Desktop\크롤링강의\2차_네이버지도_즐거찾기.py (3.7.4)" is open, showing the execution of the Selenium script. The code in the terminal includes:

```
# second_session
# 네이버 지도 즐겨찾기
from selenium import webdriver

# webdriver 설정
browser = webdriver.Chrome('./chrome/chromedriver.exe')

# 브라우저 사이즈
browser.set_window_size(1920, 1280)

# 페이지 이동
browser.get('http://www.naver.com')
```
- Browser Window:** A browser window titled "oCam (24, 21, 1804, 977)" is open, displaying a Naver map page. The page shows a search for "2차_네이버지도_즐거찾기.py" and displays a map of Seoul. The browser window also shows a "29% COMPLETE" progress bar and a "Chapter 06. Beautiful Soup" title.
- oCam Window:** A window titled "oCam (24, 21, 1804, 977)" is open, showing a video recording interface. It includes a "메뉴" (Menu) button, a "화면 녹화" (Screen Recording) button, a "게임 녹화" (Game Recording) button, and a "소리 녹음" (Sound Recording) button. The recording status is "00:00:00" and "0bytes / 31.6GB".

The bottom of the image shows a terminal window with the text "Press any key to close the terminal." and a small "eng Coding" logo in the bottom right corner.

어느 날 문짚왕 은진띠가
부탁을 해옵니다

부탁이 있어요



문짚왕 은진띠

*네이버 영화 평점과 리뷰를
저장해 두고 싶당*

통합검색

이미지

블로그

뉴스

지식iN

웹사이트

동영상

실시간검색

더보기

검색옵션

연관검색어

기억의 밤 결말

기억의 밤 줄거리

김무열

강하늘

기억의 밤 손익분기점

신고

사라진밤

기억의 밤 원작

7년의 밤

기억의 밤 나무위키

영화 기억의밤

네이버 영화

다른 사이트를 보시려면 여기를 누르세요.

기억의 밤 (Forgotten, 2017)

관람객

★★★★★

 8.41 (2,019) | 기자 평론가

★★★★★

 5.67 (6)

미스터리, 스릴러 | 2017.11.29. 개봉 | 109분 | 한국 | 15세 관람가

감독 장항준

관객수 1,387,779명

내용 새 집으로 이사 온 날 밤 정체불명의 괴한들에게 납치된 형 유석.... 더보기

다운로드

3,260

출연

리뷰

포토/동영상

AiTEMS 추천영화

편성표

이 영화에 대한 전체 407개 리뷰가 등록되어있습니다.

리뷰쓰기

기대하고 봤는데.. ci_v**** | 110 추천 | 2017.11.30.

제발 잔인하고 끔찍한 영화는 만들지 말자, 토나온다. sang**** | 71 추천 | 2017.12.02.

최고라는 말이 아깝지 않는, 강하늘과 김무열의 최고의 연기, 최고의... pjwo**** | 43 추천 | 2018.01.27.



이 부분 말씀하시는 건가요?

엄펍 쓰앵님

문짚왕 은진띠
부탁이 있어요



이왕이면
검색두 하구 싶구요..

movie.naver.com/movie/bi/mi/point.nhn?code=163788#tab

NAVER 영화

영화홈

상영작 · 예정작

- ▶ 현재 상영영화
- ▶ 개봉 예정영화
- ▶ 예고편

영화랭킹

예매

평점 · 리뷰

다운로드

인디극장 up

보기옵션

기본보기 넓게보기

알라딘

Aladdin, 2019 참여 8,049명

관람객 ★★★★★ 9.43 기자·평론가 ★★★★★ 6.00

네티즌 ★★★★★ 9.39 내 평점 ★★★★★ 등록

개요 모험, 가족, 판타지, 뮤지컬, 멜로/로맨스 | 미국 | 128분
2019.05.23 개봉

감독 가이 리치

출연 메나 마수드(알라딘), 윌 스미스(지나), 나오미 스콧(자스민) 더보기

등급 [국내] 전체 관람가 [해외] PG

총행 누적관객 12,550,932명 (09.30 기준)

성별·나이별 관람추이

41% 59%

14% 35% 31% 19%

10대 20대 30대 40대 ↑

남자 여자

20,376

네이버 영화 평점 / 리뷰

문썬왕 은진띠
부탁이 있어요



공감 제일 많이 받은 리뷰
5개 정도 봤으면 좋겠어요...☆

네티즌·관람객 평점 ▶ ★★★★★ 9.39 참여 26,542명	기자·평론가 평점 ▶ ★★★★☆ 6.00 참여 7명	내 평점 등록하기 ▶ 등록한 평점이 없습니다. 등록하기
한줄평 총 26,274건		
★★★★★ 10	월 스미스가 하드캐리하여 의도치 않게 원작을 뛰어넘은 실사영화 시민의적들(son2****) 2019.05.23 11:26 신고	공감 10423 비공감 432
★★★★★ 10	영화가 너무 아름답다..ost도 잊혀지지 않고 디즈니 실사중에서 최고중에 최고다 자스민 너무 멋지고 팬 엄청나게 생길듯 웃다가 울다가 너무 좋았음 별이(konf****) 2019.05.23 09:11 신고	공감 8044 비공감 290
★★★★★ 10	존잼임... 난 처음 도입부에 아~라~비안 나~잇 이 부분에서부터 감동받아서 울 뻔했음... 영원히고통받는롯데팬(ieie****) 2019.05.23 21:57 신고	공감 6689 비공감 247
★★★★★ 10	자스민의 speechless 명장면이고 덤으로 월 스미스의 연기력 또한 매우 굿~~ 장구(makj****) 2019.05.23 12:55 신고	공감 5645 비공감 194
★★★★★ 10	A whole new world 나올때때 개소름... 야끼우동(snst****) 2019.05.23 09:06 신고	공감 5020 비공감 186

엄펄 쓰앵님



Selenium을 써보자!

네이버 영화 평점 / 리뷰

초기 셋팅, 라이브러리 import

```
# second_session
# 네이버 영화

from selenium import webdriver
import time
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from selenium.webdriver.chrome.options import Options
from bs4 import BeautifulSoup
```

사용자에게 검색어 입력 받기

```
#검색어받기
query = str(input('검색어를 입력하세요 : '))
```

input(A)

→ A를 출력하며 사용자의 입력을 받는다.

검색어를 입력하세요 :

네이버 영화 평점 / 리뷰

브라우저 창 키기

```
# webdriver 설정
```

```
browser = webdriver.Chrome('./chrome/chromedriver.exe')
```

```
# 크롬 브라우저 대기 - 사양 고려
```

```
browser.implicitly_wait(5)
```

```
# 브라우저 사이즈
```

```
browser.set_window_size(1920, 1280)
```

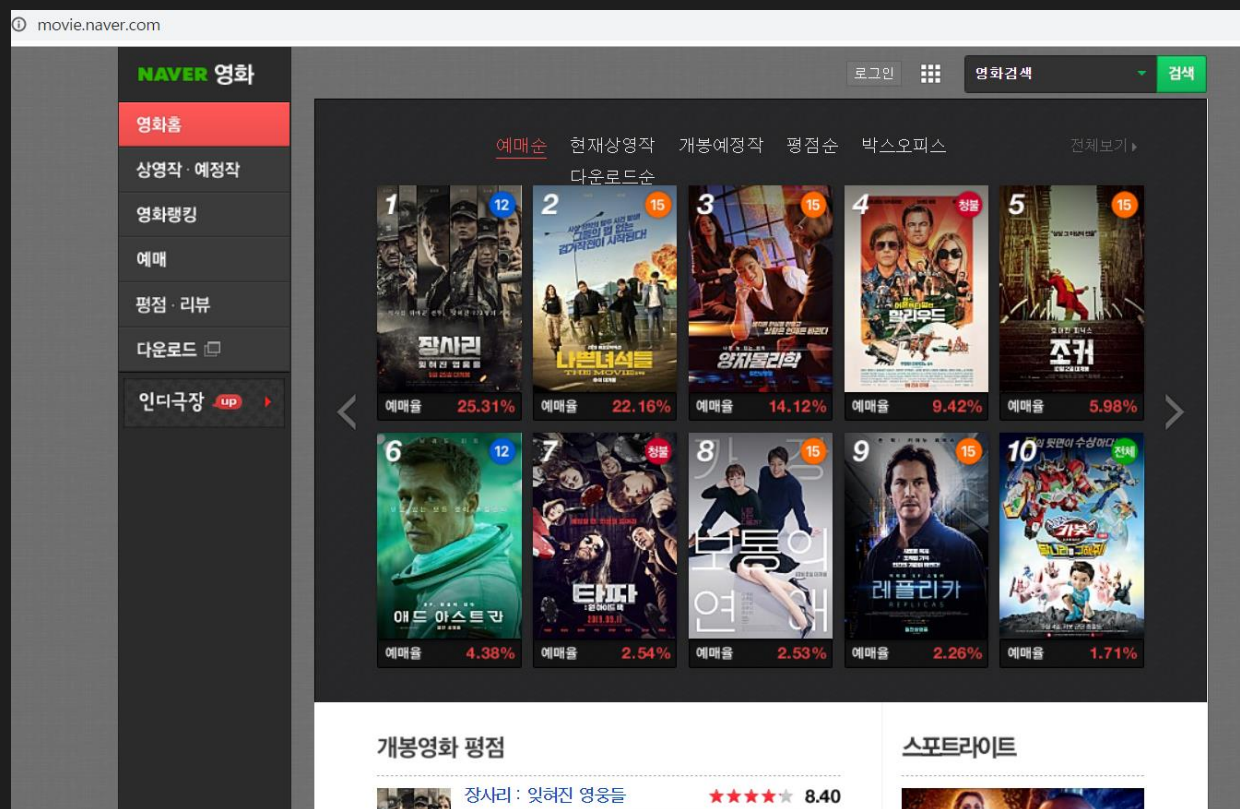
```
# 페이지 이동
```

```
browser.get('https://movie.naver.com/')
```

```
.implicitly_wait(A)
```

————→ A초 동안 대기한다.

※ 브라우저 창이 켜지나 확인!



네이버 영화 평점 / 리뷰

검색어를 입력하세요 : 알라딘

```
query = str(input('검색어를 입력하세요 : '))
```

```
# 검색창 input 선택
```

```
element = browser.find_element_by_css_selector('input#ipt_tx_srch')
```

```
# 검색어 입력
```

```
element.send_keys(query)
```

B.send_keys(A)

→ A를 B에 보낸다

※ 입력되나 확인!

The screenshot shows a web browser with a search bar at the top. The search bar is highlighted with a yellow box. A tooltip is visible over the search bar, showing the CSS selector `input#ipt_tx_srch.ipt_tx_srch` and its dimensions (147 x 29). The browser's developer tools are open, showing the HTML structure of the search form. The HTML structure includes a `<form>` tag with `id="jSearchForm"` and `action="/movie/search/resu"`, and a `<fieldset>` tag with `id="jSearchArea"` and `class="srch_area"`. Inside the `<fieldset>` tag, there is a `<div>` with `class="srch_field_on_view"` containing a `` with `class="ipt_srch"` and an `<input>` with `type="text"`, `id="ipt_tx_srch"`, `class="ipt_tx_srch"`, `name="query"`, `maxlength="100"`, `accesskey="s"`, `style="ime-mode:active;"`, and `autocomplete="off"`.

네이버 영화 평점 / 리뷰

검색클릭

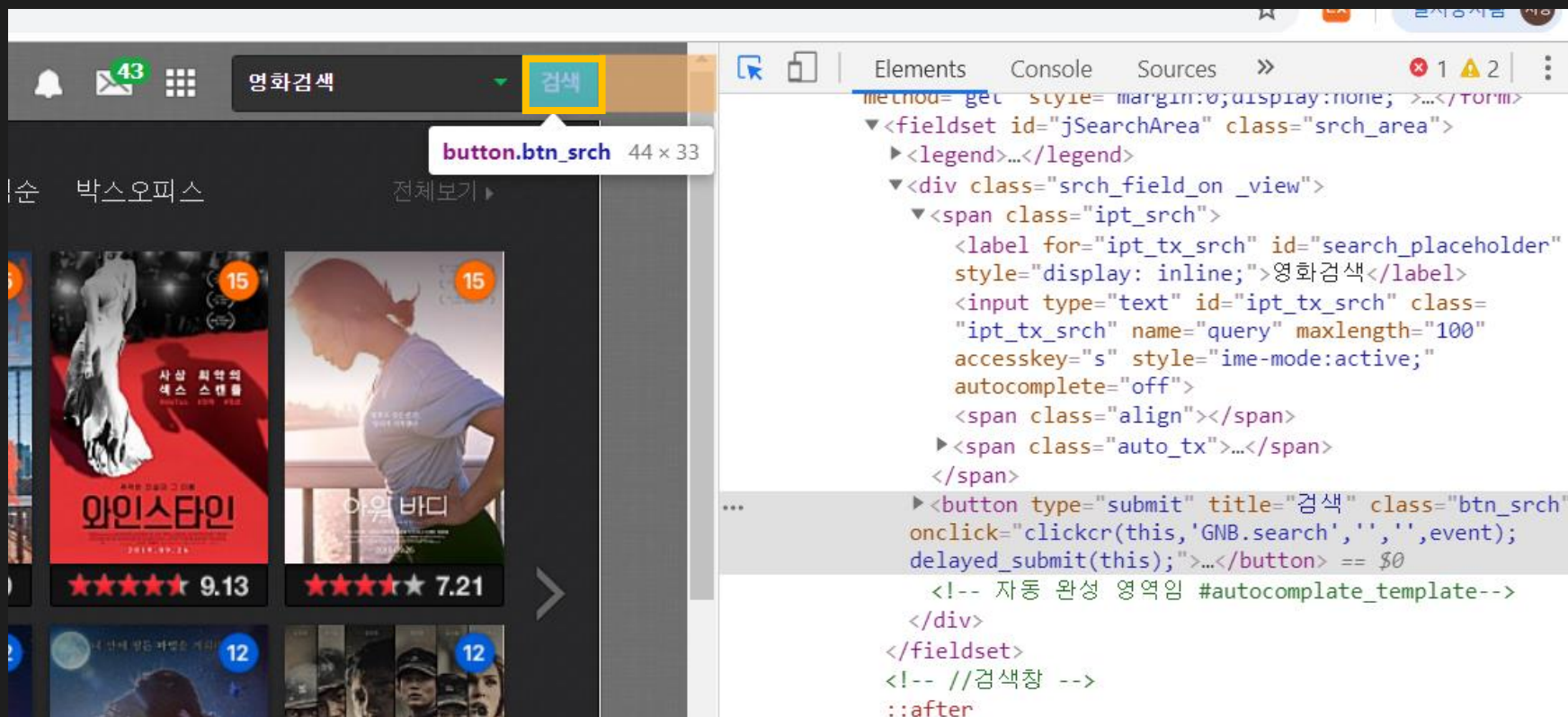
```
WebDriverWait(browser, 5).until(EC.element_to_be_clickable((By.CSS_SELECTOR, 'button.btn_srch'))).click()
```

#기다려주기

```
time.sleep(3)
```

```
WebdriverWait(A, B).until(EC.~((By. ~ ,C))).click()
```

→ A브라우저에서 C가 클릭 가능할 때까지 최대 B초 동안 기다리고, 클릭 가능해지면 클릭한다.



The screenshot shows a web browser with the Naver movie search page. The search button is highlighted with a yellow box. A tooltip shows 'button.btn_srch 44 x 33'. The browser's developer tools are open, showing the HTML structure of the search area. The HTML code in the developer tools is as follows:

```
<div class="srch_field_on _view">
  <span class="ipt_srch">
    <label for="ipt_tx_srch" id="search_placeholder"
      style="display: inline;">영화검색</label>
    <input type="text" id="ipt_tx_srch" class=
      "ipt_tx_srch" name="query" maxlength="100"
      accesskey="s" style="ime-mode:active;"
      autocomplete="off">
    <span class="align"></span>
    <span class="auto_tx">...</span>
  </span>
  <button type="submit" title="검색" class="btn_srch"
    onclick="clickcr(this,'GNB.search','',',',event);
    delayed_submit(this);">...</button> == $0
  <!-- 자동 완성 영역임 #autocomplete_template-->
</div>
</fieldset>
<!-- //검색창 -->
::after
```

네이버 영화 평점 / 리뷰

첫번째 영화 클릭

```
WebDriverWait(browser, 5).until(EC.element_to_be_clickable((By.CSS_SELECTOR, ' ? '))).click()
```



네이버 영화 평점 / 리뷰

첫번째 영화 클릭

```
WebDriverWait(browser, 5).until(EC.element_to_be_clickable((By.CSS_SELECTOR, 'li:nth-child(1) > dl > dt > a'))).click()
```



네이버 영화 평점 / 리뷰

```
# soup 초기화
```

```
soup = BeautifulSoup(browser.page_source, "html.parser")
```

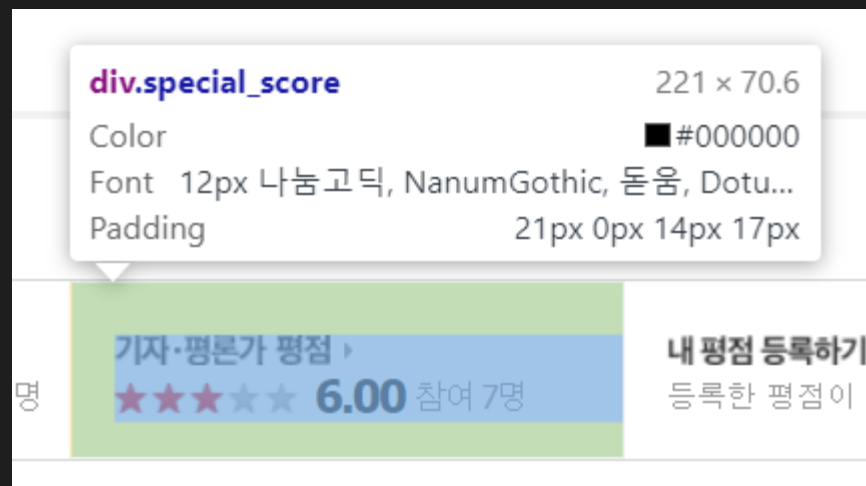
```
# 평점 출력
```

```
net_score = soup.select_one('div.netizen_score').text.strip()
```

```
print('네티즌, 관람객 평점 : ' + net_score)
```

```
spe_score = soup.select_one('div.special_score').text.strip()
```

```
print('기자, 평론가 평점 : ' + spe_score)
```



```
</div>
▼<div class="special_score">
  ▶<a href="./point.nhn?code=163"
  "off">...</a> == $0
  <!-- N=a:mra.ctitle -->
  ▼<div class="sc_view">
    ▼<div class="star_score">
      ▶<span class="st_off">...</s
      <em>6.00</em>
```

검색어를 입력하세요 : 알라딘

네티즌, 관람객 평점 : 9.39

기자, 평론가 평점 : 6.00

네이버 영화 평점 / 리뷰

```
# soup 초기화
```

```
soup = BeautifulSoup(browser.page_source, "html.parser")
```

```
# 평점 출력
```

```
net_score = soup.select_one('div.netizen_score div.sc_view em').text.strip()
```

```
print('네티즌, 관람객 평점 : ' + net_score)
```

```
spe_score = soup.select_one('div.special_score div.sc_view em').text.strip()
```

```
print('기자, 평론가 평점 : ' + spe_score)
```

div.netizen_score220 × 70.6

Color■ #000000

Font 12px 나눔고딕, NanumGothic, 돋움, Dotu...

Padding21px 0px 14px 17px

네티즌·관람객 평점 ▶

★★★★★ 9.39 참여 26,541명

기자·평론가 평점 ▶

★★★★★

div.special_score221 × 70.6

Color■ #000000

Font 12px 나눔고딕, NanumGothic, 돋움, Dotu...

Padding21px 0px 14px 17px

기자·평론가 평점 ▶

★★★★★ 6.00 참여 7명

내 평점 등록하기

등록한 평점이

```
</div>
▼<div class="special_score">
  ▶<a href="./point.nhn?code=163"
  "off">...</a> == $0
  <!-- N=a:mra.ctitle -->
  ▼<div class="sc_view">
    ▼<div class="star_score">
      ▶<span class="st_off">...</s
      <em>6.00</em>
```

```
검색어를 입력하세요 : 알라딘
네티즌, 관람객 평점 : 9.39
기자, 평론가 평점 : 6.00
```

네이버 영화 평점 / 리뷰

#리뷰 출력

```
review_list = soup.select('div[class="score_result"]')
print('<리뷰>')
for i, review in enumerate(review_list, 1):
    reple = review.select_one('p').text.strip()
    print(str(i) + '. ' + reple)
browser.quit()
```

한줄평 | 총 26,274건

★★★★★ 10

월 스미스가 하드캐리하여 의도치 않게 원작을 뛰어넘은 실사영화

시민의적들(son2****) | 2019.05.23 11:26 | 신고

공감 10423 비공감 432

★★★★★ 10

영화가 너무 아름답다..ost도 잊혀지지 않고 디즈니 실사중에서 최고중에 최고다 자스민 너무 멋지고 팬 엄청나게 생길듯 웃다가 울다가 너무 좋았음

별이(konf****) | 2019.05.23 09:11 | 신고

공감 8044 비공감 290

★★★★★ 10

존잼임... 난 처음 도입부에 아~라~비안 나~잇 이 부분에서부터 감동받아서 울 뻔했음...

```
<div class="score_result">
```

```
<ul>
```

```
<li>
```

```
<div class="star_score">...</div>
```

```
<div class="score_reple">
```

```
<p>월 스미스가 하드캐리하여 의도치 않게 원  
넘은 실사영화 </p> == $0
```

```
<dl>...</dl>
```

```
</div>
```

네이버 영화 평점 / 리뷰

#리뷰 출력

```
review_list = soup.select('div.obj_section:nth-child(5) div.score_reple')
print('<리뷰>')
for i, review in enumerate(review_list, 1):
    reple = review.select_one('p').text.strip()
    print(str(i) + '. ' + reple)
browser.quit()
```

한줄평 | 총 26,274건

★★★★★ 10 월 스미스가 하드캐리하여 의도치 않게 원작을 뛰어넘은 실사영화

시민의적들(son2****) | 2019.05.23 11:26 | 신고

공감 10423 비공감 432

★★★★★ 10 영화가 너무 아름답다..ost도 잊혀지지 않고 디즈니 실사중에서 최고중에 최고다 자스민 너무 멋지고 팬 엄청나게 생길듯 웃다가 울다가 너무 좋았음

별이(konf****) | 2019.05.23 09:11 | 신고

공감 8044 비공감 290

★★★★★ 10 존잼임... 난 처음 도입부에 아~라~비안 나~잇 이 부분에서부터 감동받아서 울 뻔했음...

```
<div class="score_result">
  <ul>
    <li>
      <div class="star_score">...</div>
      <div class="score_reple">
        <p>월 스미스가 하드캐리하여 의도치 않게 원
        넘은 실사영화 </p> == $0
      </div>
    </li>
  </ul>
</div>
```

#브라우저 종료

browser.quit()



검색어를 입력하세요 : 알라딘

네티즌, 관람객 평점 : 9.39

기자, 평론가 평점 : 6.00

<리뷰>

1. 윌 스미스가 하드캐리하여 의도치 않게 원작을 뛰어넘은 실사영화

2. 영화가 너무 아름답다..ost도 잊혀지지 않고 디즈니 실사중에서 최고중에 최고다

자스민 너무 멋지고 팬 엄청나게 생길듯 웃다가 울다가 너무 좋았음

3. 존쟁임... 난 처음 도입부에 아~라~비안 나~잇 이 부분에서부터 감동받아서 울 뻔했음...

4. 자스민의 speechless 명장면이고 덤으로 윌 스미스의 연기력 또한 매우 굿~~

5. A whole new world 나올때때 개소름...

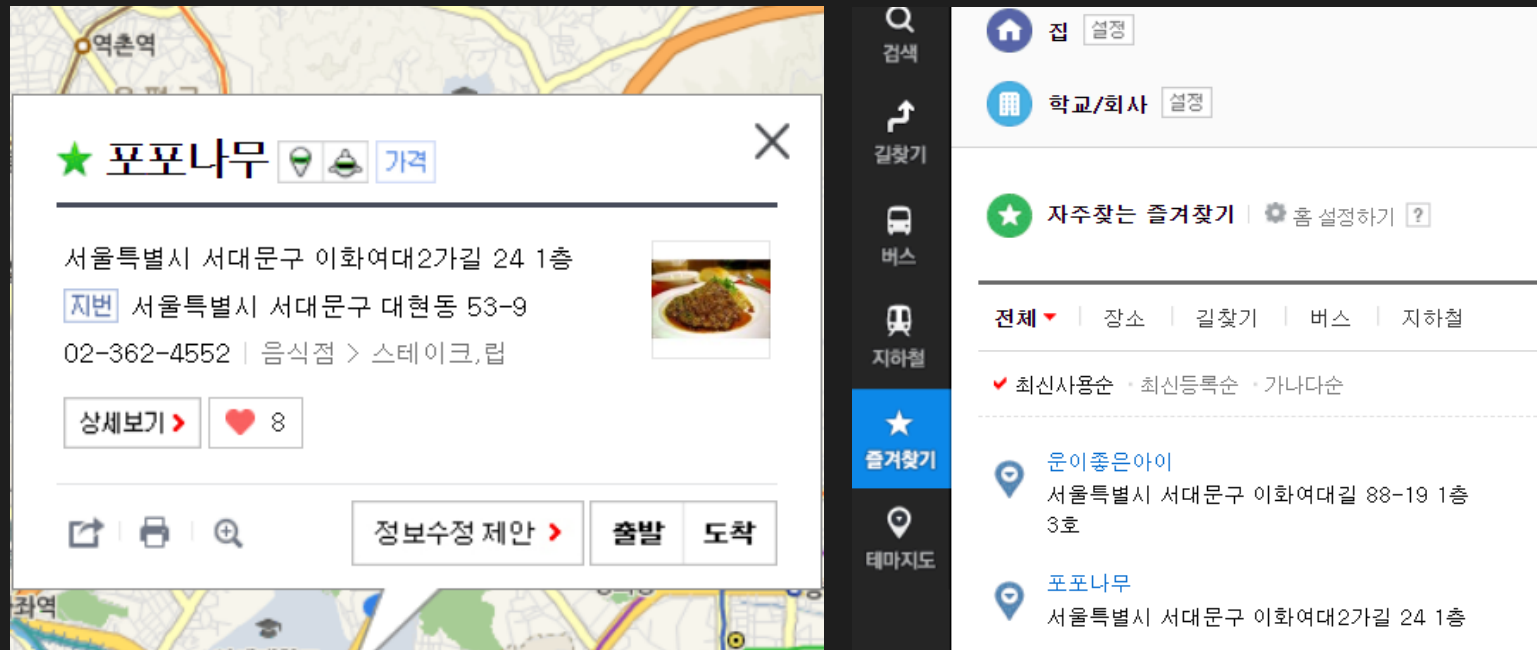
네이버 지도 즐겨찾기

새 파일 만든 후 주석 달기

```
*2차_네이버지도.py - C:\Users\user\Desktop\워크롤
File Edit Format Run Options Window Help
# second_session
# 네이버 지도
```

코드 적을 곳

네이버 로그인 후 즐겨 찾기 몇 개 해두기



네이버 지도 즐겨찾기

초기 셋팅, 라이브러리 import

```
from selenium import webdriver
import time
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from selenium.webdriver.chrome.options import Options
from bs4 import BeautifulSoup
```

webdriver 설정

```
browser = webdriver.Chrome('./chrome/chromedriver.exe')
```

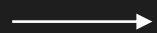
크롬 브라우저 대기 - 사양 고려

```
browser.implicitly_wait(5)
```

브라우저 사이즈

```
browser.set_window_size(1920, 1280)
```

`time.sleep(A)`



A초 동안 대기한다.

페이지 이동

```
browser.get('https://nid.naver.com/')
time.sleep(3)
```

네이버 지도 즐겨찾기

로그인 하기 - 아이디/비밀번호 입력

아이디입력

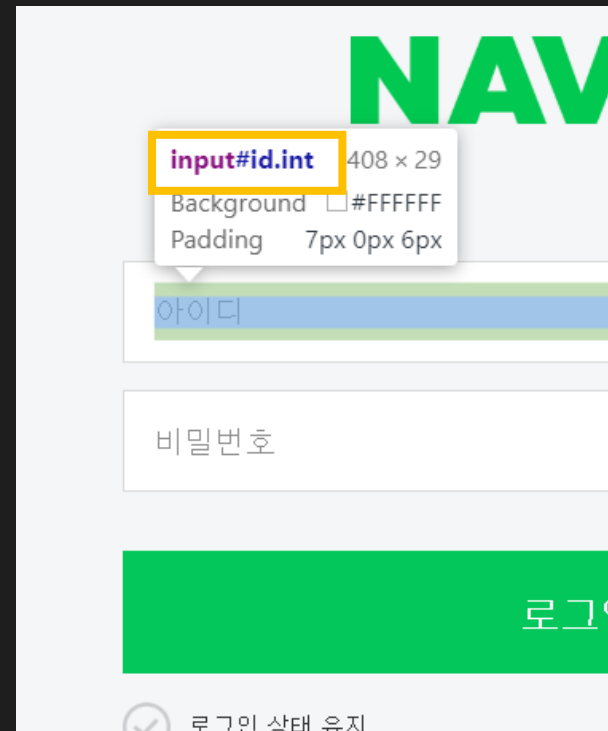
```
id_element = browser.find_element_by_css_selector('input#id.int')  
id_element.send_keys('tjdud0123')
```

패스워드입력

```
pw_element = browser.find_element_by_css_selector('input#pw.int')  
pw_element.send_keys('비번')
```

로그인클릭

```
WebDriverWait(browser, 5).until(EC.element_to_be_clickable((By.CSS_SELECTOR, 'input#login.int'))).click()
```



네이버 지도 즐겨찾기

로그인 하기 - 아이디/비밀번호 입력

아이디입력

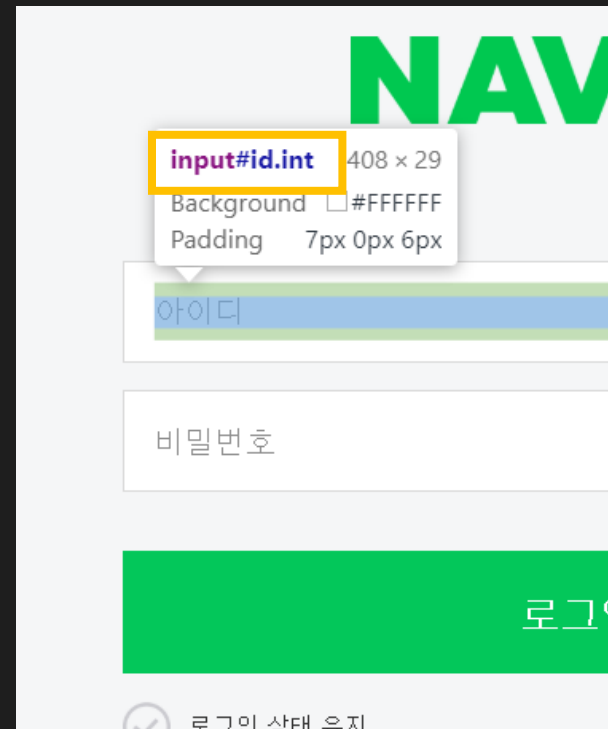
```
id_element = browser.find_element_by_css_selector('input#id.int')
id_element.send_keys('tjdud0123')
```

패스워드입력

```
pw_element = browser.find_element_by_css_selector('input#pw.int')
pw_element.send_keys('비번')
```

로그인클릭

```
WebDriverWait(browser, 5).until(EC.element_to_be_clickable((By.CSS_SELECTOR, 'input.btn_global'))).click()
```



네이버 지도 즐겨찾기

아래 이미지를 보이는 대로 입력해주세요.



새로고침

음성으로 듣기

자동입력 방지문자

===== RESTART: C:\Users\swu

=====

자동방지 문자입력 : 12B5E387

자동방지입력, 패스워드입력

```
time.sleep(5)
```

```
pw_element = browser.find_element_by_css_selector('input#pw.int')
```

```
pw_element.send_keys('dja2533')
```

```
time.sleep(5)
```

```
auto_element = browser.find_element_by_css_selector('input#chptcha.int')
```

```
string = input('자동방지 문자입력 : ')
```

```
auto_element.send_keys(string)
```

```
WebDriverWait(browser, 5).until(EC.element_to_be_clickable((By.CSS_SELECTOR, 'input.btn_global'))).click()
```

네이버 지도 즐겨찾기

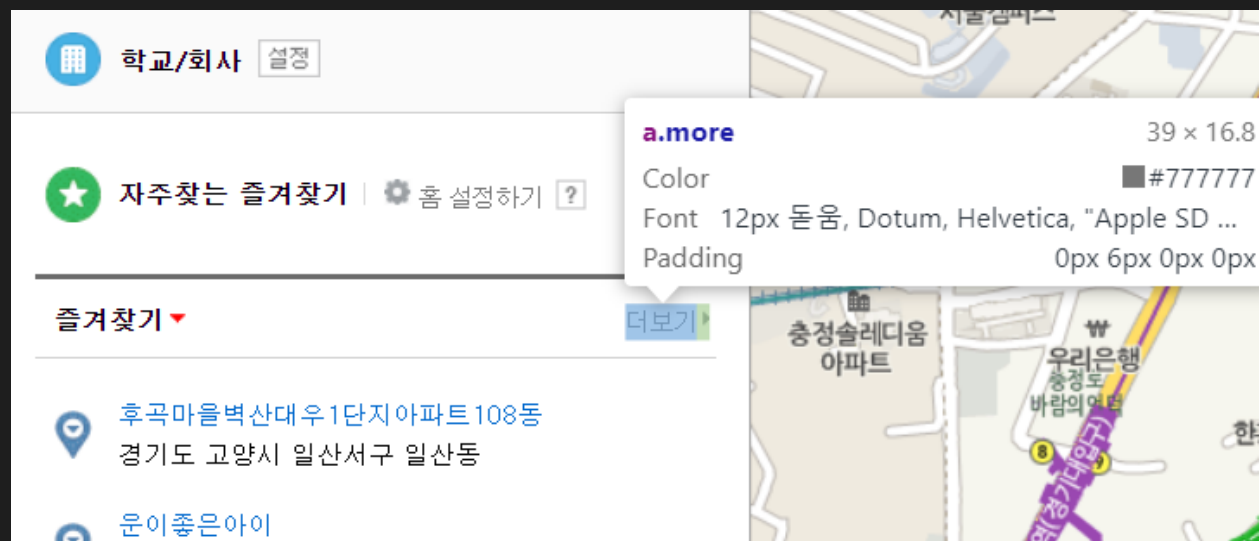
```
# 페이지 이동 후 즐겨찾기 클릭
```

```
time.sleep(5)
```

```
browser.get('https://map.naver.com/')
```

```
time.sleep(5)
```

```
WebDriverWait(browser, 5).until(EC.element_to_be_clickable((By.CSS_SELECTOR, 'a.more'))).click()
```



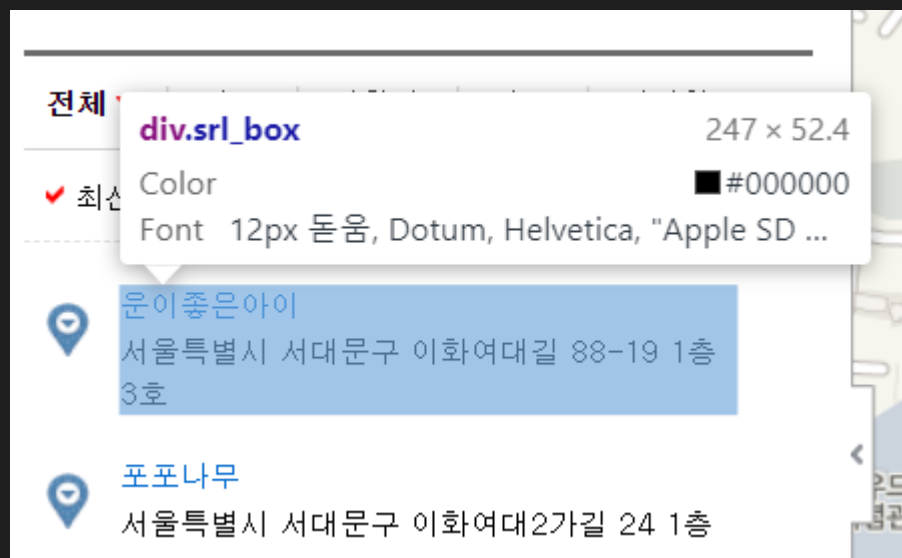
네이버 지도 즐겨찾기

#골라내기

```
soup = BeautifulSoup(browser.page_source, "html.parser")
like_list = soup.select('div.srl_box')
```

```
print('<내 즐겨찾기>')
```

```
for i, like in enumerate(like_list, 1):
    name = like.select_one('div.srl_box').text.strip()
    location = like.select_one('div.srl_box').text.strip()
    print(name)
    print(location)
    print('-----')
```



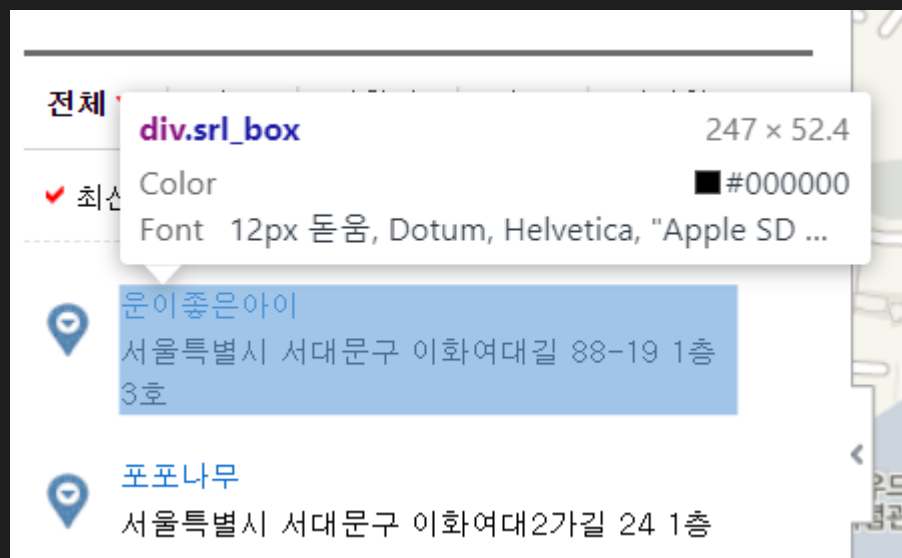
네이버 지도 즐겨찾기

#골라내기

```
soup = BeautifulSoup(browser.page_source, "html.parser")
like_list = soup.select('div.srl_box')
```

```
print('<내 즐겨찾기>')
```

```
for i, like in enumerate(like_list, 1):
    name = like.select_one('a.srl_a').text.strip()
    location = like.select_one('div.srl_d').text.strip()
    print(name)
    print(location)
    print('-----')
```



네이버 지도 즐겨찾기

#브라우저 종료
browser.quit()



운이좋은아이
서울특별시 서대문구 이화여대길 88-19 1층 3호

포포나무
서울특별시 서대문구 이화여대2가길 24 1층

남구국민체육센터
부산 남구 백운포로 110 남구국민체육센터

고양스파랜드사우나
경기 고양시 덕양구 중양로 554

IKEA 고양점
경기 고양시 덕양구 권율대로 420

경기도 고양시 덕양구 화정동 925-4 주소

후곡마을벽산대우1단지아파트108동
경기도 고양시 일산서구 일산동

현대파리지앵 주차장
서울 양천구 목동서로 57 현대파리지앵

슈라즈케이크
경기 고양시 일산서구 호수로856번길 34-4

웁스 해운대점
부산 해운대구 중동1로 31

한빛마을센트럴파크한라비발디아파트109동

<내 즐겨찾기>
운이좋은아이
서울특별시 서대문구 이화여대길 88-19 1층

포포나무
서울특별시 서대문구 이화여대2가길 24 1층

남구국민체육센터
부산 남구 백운포로 110 남구국민체육센터

고양스파랜드사우나
경기 고양시 덕양구 중양로 554

IKEA 고양점
경기 고양시 덕양구 권율대로 420

경기도 고양시 덕양구 화정동 925-4 주소

후곡마을벽산대우1단지아파트108동
경기도 고양시 일산서구 일산동

현대파리지앵 주차장
서울 양천구 목동서로 57 현대파리지앵

슈라즈케이크
경기 고양시 일산서구 호수로856번길 34-4

웁스 해운대점
부산 해운대구 중동1로 31

한빛마을센트럴파크한라비발디아파트109동
경기도 파주시 야왕동

뉴욕11번가엔치킨클럽
서울 서대문구 연세로11길 35

엠티993
부산 해운대구 달맞이길62번길 78

좋아하는 웹툰 최근 10화 베댓 긁어 오기

NAVER 만화 | 웹소설

제목/작가로 검색할 수 있습니다.

Q

홈

웹툰

베스트 도전

도전만화

마이페이지

단행본만화

장르소설

요일별

장르별

작품별

작가별

연도별

테마웹툰

완결웹툰

요일전체

월요일웹툰

화요일웹툰

수요일웹툰

목요일웹툰

금요일웹툰

토요일웹툰

일요일웹툰

유일무이 로맨스

두부

평범한 최준생 공유일의 일상에 인기절정 탑배우 탁무이와 아이돌 류민이 침범해오기 시작했다!

스토리, 일상, 드라마, 로맨스 | 전체연령가

⊕ 관심웹툰

첫화보기

작가의 다른 작품

99,999+

이미지	제목	별점
	다음 화를 미리 만나보세요.	
	77화	★★★★★ 9.99
	76화	★★★★★ 9.99
	75화	★★★★★ 9.99
	74화	★★★★★ 9.99
	73화	★★★★★ 9.99

77화 베댓 추출
-미친거 아냐??? 진짜면 진짜 도른년이고 연기면 제배역 찾았네
-...?? 아니 저건 연기가 아니고 자기 자신이 무이 머릿속에 완전 제대로 박혔으면 해서 일부러 그런거잖아; 미친거 아닌가 저건 진짜 욕 존나 나오는데..?? 제가 저 사람 경찰서에 넣어도 되나요 진짜 미쳤다
-꿈일것같은데?라면 다먹고 잠든게 아닐까?
-제발 다음화 제발 다음화 ㅠㅠㅠㅠㅠ다음화 주세요
-다음화가 필요해...
-베댓 카페 근무잔데 저런 아메리카노 물 95도 넘어요... ㅠㅠㅠ 겁나 뜨거운데 화상 입은 손님도 있었어요..... 저거 저러면 큰일남...ㅠㅠ
-제발 상상이어라... 제발...
-이디야 레시피 스팀우유온도 55도-60도에서 멈추게되었습니다잔열로 80-90도까지 올라가기때문이죠 아메리카노나 차종류에 쓰는 물은 90도로 아예 맞춰져있습니다다저렇게 부어버리면 화상 당연히입습니다
-안녕하세요 현직 바리스타입니다. 머신기에서 나오는 물은 보통 90도 이상입니다. 피부가 약하면 겁나 따갑고 아파요. 아니 근데 머신기에서 저렇게 물을 뽑으면 머신기 상할텐데 정수기없는거야? 예반데 아무튼 존나 뜨겁습니다. 90도 이상이 어느정도냐면요. 잠깐 담궜도 빨갈게 부어오를정도에요...카페서 일하는 사람은 항상 화상에 손이 데이고 칼에 베이죠 ㅎㅎㅎ... 힘쇼...그리고 무이야 빨리 병원가... 아 여기서 결제하는게 아니었다... 한주만 더 존버할걸...
-뭔소리야 베댓.. 나일했던카페는 항상 85도 이상이었는데
-잇몸 마르게 웃으면서 보다가 개놀람 뭐야 ㅁㅆ...
-연기라고 해도 말이 안되고... 실체라고 해도 말이 안되는데... 연기라면 어느 작가가 저런 시나리오를 쓰고, 실체라면 저렇게 많은 사람들 앞에서 저렇게 할수 있을까?
-아...!바...깜박이좀 켜고 들어와라 도ㄹ이녀나 ㅠㅠㅠㅠㅠㅠㅠㅠㅠㅠ
-전직 바리스타입니다.. 작년에 딱 저렇게 물 담긴 컵이 손 전체에 옆질러졌다가 두달 이상을 붓대감고다녔습니다.. 일 당연히 못했도 관리 진짜 잘했음에도 흉터도 조금남았어요ㅠㅠㅠ 최소 2도화상 입어요..... 90도 이상되는 물 상상안되지겠지만 진짜 한방울만 튀어도 피부약한분은 바로 물집생깁니다당ㅠ영영 막팔보자마자 제 손에 옆질러졌을때 생각나서 소름돋았어요 ㅠ 무이는 다치지 않았길 제-발 !

76화 베댓 추출
-앵 의사친구도 연애인 출신이엇어와 무2친구 엄친이였네
-기자가 의외로 활약해서 스토커나 잡아줬으면...ㅎ
-영수 퇴원기념파티가 탁무이 사심채우기가 되어버렸어 ㅋㅋㅋ
-헐 설마 소고기넣고 끓인 라면...?
-탁무이 친구 연예인이었던거 몰랐던 사람 많았구나 ...ㅋㅋ 고당때 탁무이 화상장면에서도 같이 촬영하고 그러는거 나왔었는데
-28화에 의사친구 연예인였다는거 나와여영수 다홍 의사친구랑 다 같이 아역으로 출연

네이버 웹툰 베댓

새 파일 만든 후 주석 달기

```
2차_네이버웹툰.py - C:\Users\User\Desktop\크롤
File Edit Format Run Options Window Help
# second_session
# 네이버 웹툰 베댓
```

초기 셋팅, 라이브러리 import

```
from selenium import webdriver
import time
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from selenium.webdriver.chrome.options import Options
from bs4 import BeautifulSoup
import urllib.parse as rep
import urllib.request as req
from bs4 import BeautifulSoup
from fake_useragent import UserAgent
```

Webdriver 설정

```
# webdriver 설정
browser = webdriver.Chrome('./chrome/chromedriver.exe')

# 크롬 브라우저 대기 - 사양 고려
browser.implicitly_wait(5)

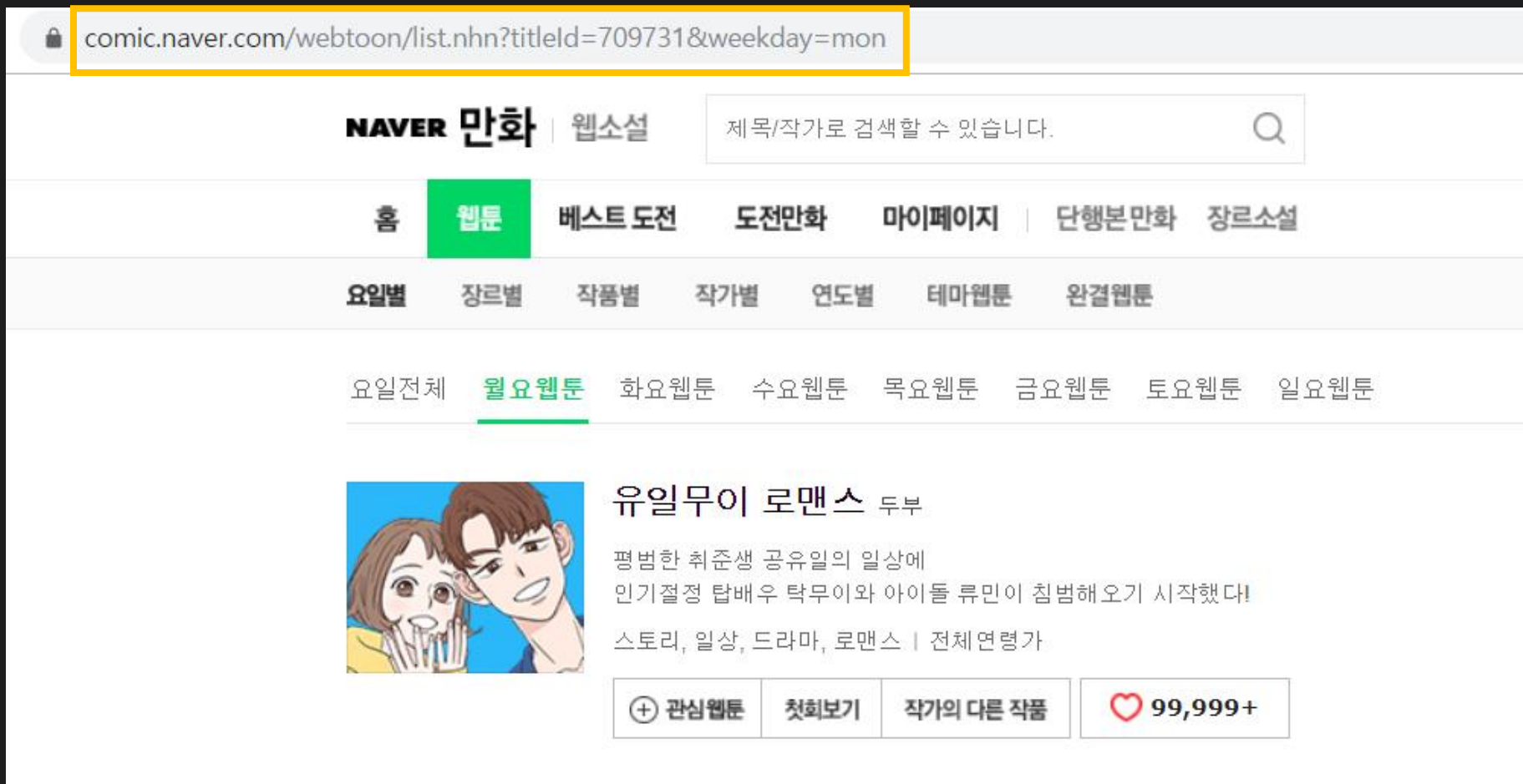
# 브라우저 사이즈
browser.set_window_size(1920, 1280)
```


네이버 웹툰 베댓

페이지 이동

```
my_webtoon = 'https://comic.naver.com/webtoon/list.nhn?titleId=695796&weekday=sun'  
browser.get(my_webtoon)  
time.sleep(3)
```

원하는 웹툰 url 저장!



네이버 웹툰 배댓

링크 저장

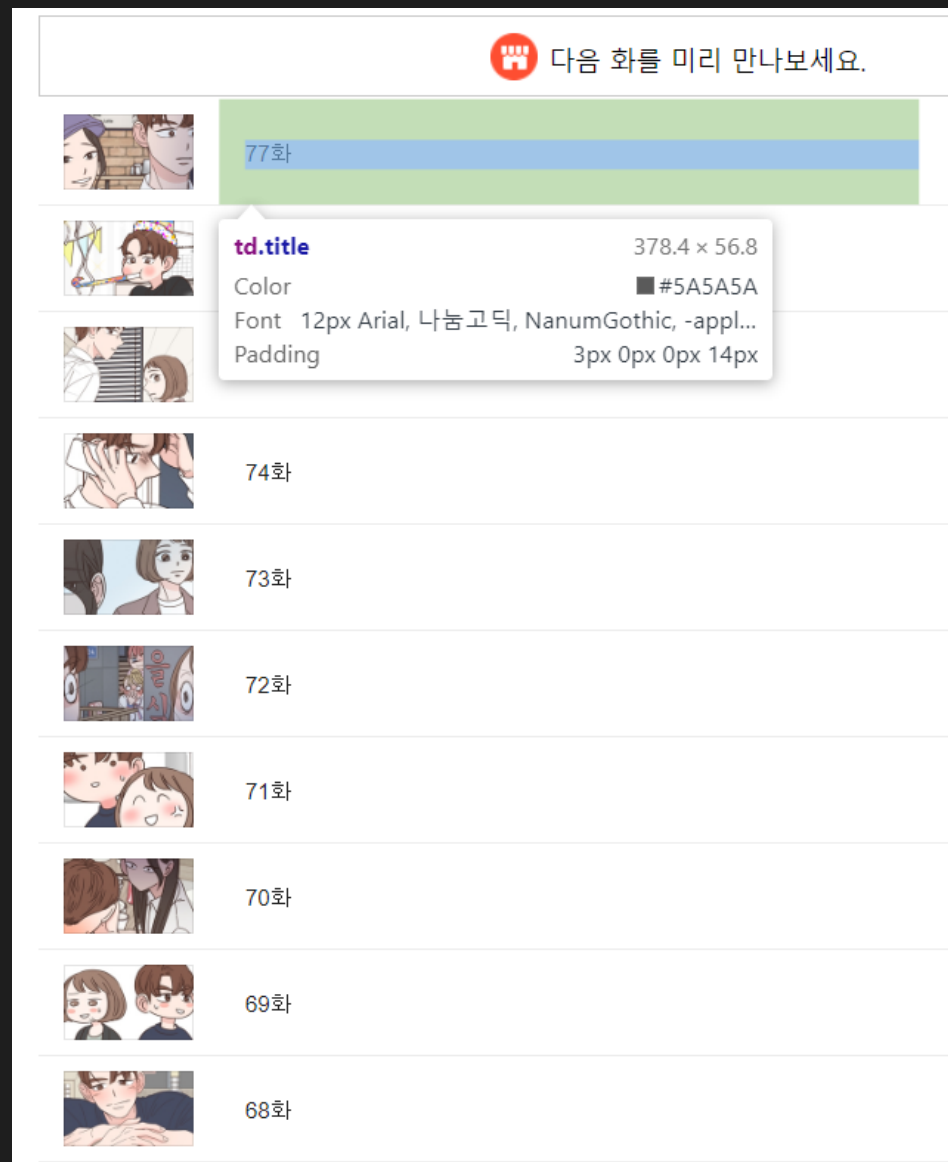
```
baseurl = 'https://comic.naver.com'
soup = BeautifulSoup(browser.page_source, "html.parser")
links = soup.select("td.title > a")
next_url=[]
for link in links:
    next_url.append(baseurl + link['href'])
```

A.append(B)

→ B를 A에 저장한다.

```
<tr>
  > <td>...</td>
  ▼ <td class="title"> == $0
    <a href="/webtoon/detail.nhn?
      titleId=709731&no=77&weekday=mon" onclick=
        "nclk_v2(event,'lst.title','709731','77
      </a>
```

※ print(next_url) 해보기!





관찰



생각



질문

네이버 웹툰 베댓

#각 화에 들어가서 댓글들 추출

for url in next_url:

```
# 몇 화인지 프린트
```

```
start_index = int(url.index('no')) + 3
```

```
last_index = int(url.rindex('&'))
```

```
print(url[start_index : last_index] + '화 베댓 추출')
```

```
#링크 이동 후 댓글 프레임 가져오기
```

```
browser.get(url)
```

```
time.sleep(5)
```

```
browser.switch_to.frame('commentIframe')
```

```
soup = BeautifulSoup(browser.page_source, "html.parser")
```

```
#댓글들 추출, 댓글출력call
```

```
reples = soup.select('div.u_cbox_text_wrap')
```

```
print_reple(reples)
```

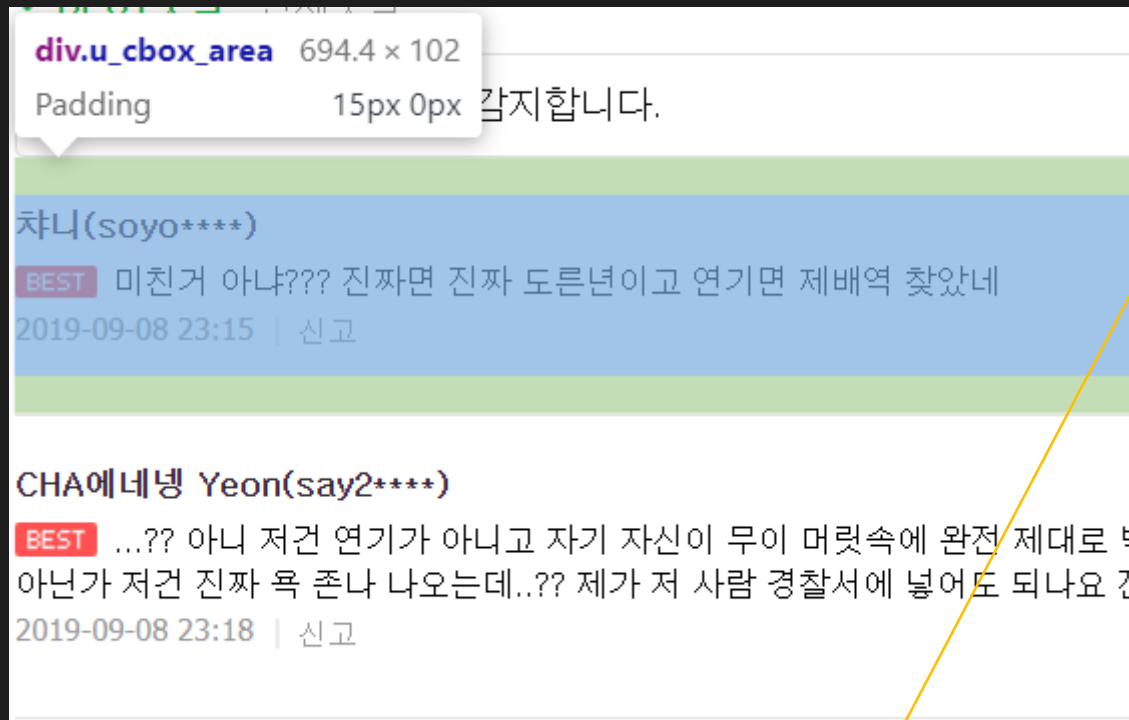
```
print('-----')
```



시간상
따로 질문!

네이버 웹툰 베타

soup.select('div.u_cbox_area')?!



※ iframe : 새로운 html 파일영역!

```
<iframe id="commentIframe" width="100%"
frameborder="0" scrolling="no" title="유일무이 로만
- 77화 댓글영역" style="min-height: 500px; height:
1946px;" src="/comment/comment.nhn?
titleId=709731&no=77">
#document
<!doctype html>
<html lang="ko">
  <head>...</head>
  <body>
    <div id="cbox_module" class="u_cbox">
      <div class="u_cbox_wrap u_cbox_ko
u_cbox_type_sort_best">
        <div class="u_cbox_head">...</div>
        <div class="u_cbox_write_wrap">...</div>
        <div class="u_cbox_notice
u_cbox_notice_none" style="display:
none;"></div>
        <div class="u_cbox_sort">...</div>
        <div class="u_cbox_cleanbot">...</div>
        <div class="u_cbox_content_wrap">
          <ul class="u_cbox_list">
            <li class="u_cbox_comment
cbox_module__comment_376538347
_user_id_no_446XA" data-info=
"commentNo:376538347,deleted:
false,best:true,visible:true,secret
false,manager:false,mine:false,repo
undefined,blindReport:false,objectI
'709731_77',replyLevel:
1,parentCommentNo:376538347">
              <div class="u_cbox_comment_box">
                <div class="u_cbox_area">...</div>
              </div>
            <div class="u_cbox_reply_area">
```

네이버 웹툰 베댓

```
#각 화에 들어가서 댓글들 추출  
for url in next_url:
```

`.switch_to.frame(B)`

→ B라는 iframe 문서로 이동.

```
# 몇 화인지 프린트
```

```
start_index = int(url.index('no')) + 3  
last_index = int(url.rindex('&'))  
print(url[start_index : last_index] + '화 베댓 추출')
```

```
#링크 이동 후 댓글 프레임 가져오기
```

```
browser.get(url)  
time.sleep(5)  
browser.switch_to.frame('commentIframe')  
soup = BeautifulSoup(browser.page_source, "html.parser")
```

```
#댓글들 추출, 댓글출력call
```

```
replies = soup.select('div.u_cbox_text_wrap')  
print_reple(replies)  
print('-----')
```

```
<iframe id="commentIframe" width="100%"  
frameborder="0" scrolling="no" title="유일무이 로맨  
- 77화 댓글영역" style="min-height: 500px; height:  
1946px;" src="/comment/comment.nhn?  
titleId=709731&no=77">
```

```
▼#document
```

```
<!doctype html>
```

```
▼<html lang="ko">
```

```
▶<head>...</head>
```

```
▼<body>
```

```
▼<div id="cbox_module" class="u_cbox">
```

```
▼<div class="u_cbox_wrap u_cbox_ko  
u_cbox_type_sort_best">
```

```
▶<div class="u_cbox_head">...</div>
```

```
▶<div class="u_cbox_write_wrap">...</div>
```

```
<div class="u_cbox_notice  
u_cbox_notice_none" style="display:  
none;"></div>
```

```
▶<div class="u_cbox_sort">...</div>
```

```
▶<div class="u_cbox_cleanbot">...</div>
```

```
▼<div class="u_cbox_content_wrap">
```

```
▼<ul class="u_cbox_list">
```

```
▼<li class="u_cbox_comment  
cbox_module__comment_376538347  
_user_id_no_446XA" data-info=  
"commentNo:376538347,deleted:  
false,best:true,visible:true,secret  
false,manager:false,mime:false,repo  
undefined,blindReport:false,objectI  
'709731_77',replyLevel:  
1,parentCommentNo:376538347">
```

```
▼<div class="u_cbox_comment_box">
```

```
▶<div class="u_cbox_area">...</di
```

```
</div>
```

네이버 웹툰 베댓

#각 화에 들어가서 댓글들 추출

```
for url in next_url:
```

```
# 몇 화인지 프린트
```

```
start_index = int(url.index('no')) + 3
```

```
last_index = int(url.rindex('&'))
```

```
print(url[start_index : last_index] + '화 베댓 추출')
```

```
#링크 이동 후 댓글 프레임 가져오기
```

```
browser.get(url)
```

```
time.sleep(5)
```

```
browser.switch_to.frame('commentIframe')
```

```
soup = BeautifulSoup(browser.page_source, "html.parser")
```

```
#댓글들 추출, 댓글출력call
```

```
replies = soup.select('div.u_cbox_text_wrap')
```

```
print_reple(replies)
```

```
print('-----')
```

div.u_cbox_text_wrap 694.4 × 17.6

BEST 꿈일것같은데?라면 다먹고 잠든게 아닐까?

2019-09-09 01:49 | 신고

```
<div class="u_cbox_info">...</div>
```

```
<div class="u_cbox_text_wrap"> == $0
```

```
<span class="u_cbox_ico_best">BEST</span>
```

```
<span class="u_cbox_contents" data-lang="ko">꿈일것같은데?라면 다먹고 잠  
게 아닐까?</span>
```



쿼터 구

Q. 함수가 뭔가요?



수원

```
print_reple(reples)
```

Q. 괄호 안의 것을 무엇이라 하나요?

네이버 웹툰 배댓

댓글 출력

```
def print_reple(reples):  
    for reple in reples:  
        try:  
            print('-' + reple.select_one('span.u_cbox_contents').text.strip())  
        except:  
            pass
```

print_reple(reples)

```
<div class="u_cbox_info">...</div>  
<div class="u_cbox_text_wrap">== $0  
    <span class="u_cbox_ico_best">BEST  
    </span>  
    <span class="u_cbox_contents" data-  
    lang="ko">꿈일것같은데?라면 다먹고 잠  
    게 아닐까?</span>
```

노송희(ds54****)

⚠ BEST 클린봇이 이용자 보호를 위해 숨긴 댓글입니다.
2019-09-09 00:44

예외처리 필요!

차니(Keepon****)
BEST 미친거 아니야?? 진짜면 진짜 도련님이요 연기면 제배역 찾았네
2019-09-09 00:45 신고

CHA에네냉 Yeon(say2****)
BEST ...?? 아니 저건 연기가 아니라고 자기 자신이 무이 머릿속에 완전 제대로 막혔으면 해서 일부러
이런가 저런 진짜 욕 존나 나오는데...?? 제가 저 사람 경찰서에 넣어도 되나요 진짜 미쳤다
2019-09-09 00:45 신고

lemnot(sosu****)
BEST 꿈일것같은데?라면 다먹고 잠든게 아닐까?
2019-09-09 01:49 신고

반은(sagitt****)
BEST 제발 다음화 제발 다음화 뉴 뉴 뉴 다음화 주세요
2019-09-09 23:15 신고

과연화(****)
BEST 다음화가 필요해...
2019-09-09 23:15 신고

네이버 웹툰 배댓

댓글 출력

```
def print_reple(reples):  
    for reple in reples:  
        try:  
            print('-' + reple.select_one('span.u_cbox_contents').text.strip())  
        except:  
            pass
```

#각 화에 들어가서 댓글들 추출

for url in next_url:

몇 화인지 프린트

start_index = int(url.index('no')) + 3

last_index = int(url.rindex('&'))

print(url[start_index : last_index] + '화 배댓 추출')

#링크 이동 후 댓글 프레임 가져오기

browser.get(url)

time.sleep(5)

browser.switch_to.frame('commentIframe')

soup = BeautifulSoup(browser.page_source, "html.parser")

#댓글들 추출, 댓글출력call

reples = soup.select('div.u_cbox_text_wrap')

print_reple(reples)

print('-----')

※ 함수는 호출 전 선언 먼저!
순서 유의!

#브라우저 종료
browser.quit()



77화 베댓 추출

-미친거 아냐??? 진짜면 진짜 도른년이고 연기면 제배역 찾았네
-...?? 아니 저건 연기가 아니고 자기 자신이 무이 머릿속에 완전 제대로 박혔으면 해서 일부러 그런거잖아; 미친거 아냐가 저건 진짜 욕 존나 나오는데..?? 제가 저 사람 경찰서에 넣어도 되나요 진짜 미쳤다
-꿈일것같은데?라면 다먹고 잠든게 아닐까?
-제발 다음화 제발 다음화 ㅠㅠㅠㅠ 다음화 주세요
-다음화가 필요해...
-베댓 카페 근무잔데 저런 아메리카노 물 95도 넘어요... ㅠㅠㅠ 겁나 뜨거운데 화상 입은 손님도 있었어요..... 저거 저러면 큰일남...ㅠㅠ
-제발 상상이어라.. 제발..
-이디야 레시피 스팀우유온도 55도-60도에서 멈추게되어있습니다잔열로 80-90도까지 올라가기때문이죠 아메리카노나 차종류에 쓰는 물은 90도로 아예 맞춰져있습니다저렇게 무어버리면 화상 당연히입습니다
-안녕하세요 현직 바리스타입니다. 머신기에서 나오는 물은 보통 90도 이상입니다. 피부가 약하면 겁나 따갑고 아파요. 아니 근데 머신기에서 저렇게 물을 뽑으면 머신기 상할텐데 정수기없는거야? 예반데 아무튼 존나 뜨겁습니다. 90도 이상이 어느정도냐면요. 잠깐 담궜도 빨갈게 무어오를정도에요...카페서 일하는 사람은 항상 화상에 손이 데이고 칼에 베이죠 ㅎㅎ... 힘쇼...그리고 무이야 빨리 병원가... 아 여기서 결제하는게 아니었다.. 한주만 더 존버할께..
-뭔소리야 베댓.. 나일했던카페는 항상 85도 이상이었는데
-잇몸 마르게 웃으면서 보다가 개놀람 뭐야 ㅁㅁ..
-연기라고 해도 말이 안되고... 실제라고 해도 말이 안되는데... 연기라면 어느 작가가 저런 시나리오를 쓰고, 실제라면 저렇게 많은 사람들 앞에서 저렇게 할수 있을까?
-아...! 바깥이 좀 쾜고 들어와라 도르니녀나 ㅠㅠㅠㅠㅠㅠㅠㅠ
-전직 바리스타입니다.. 작년에 딱 저렇게 물 담긴 컵이 손 전체에 얹질러졌다가 두달 이상을 붓대감고다녔습니다.. 일 당연히 못해도 관리 진짜 잘했음에도 흉터도 조금남았어요ㅠㅠ 최소 2도화상 입어요..... 90도 이상되는 물 상상안드시겠지만 진짜 한방울만 튀어도 피부약한분은 바로 물집생깁니다당영영 막팔보자마자 제 손에 얹질러졌을때 생각나서 소름돋았어요 ㅏ 무이는 다치지 않았길 제-발!

76화 베댓 추출

-영 의사친구도 연애인 출신이었어와 무2친구 엄친이었네
-기자가 의외로 활약해서 스토크나 잡아줬으면...ㅎ
-영수 퇴원기념파티가 탁무이 사심채우기가 되어버렸어 ㄹㄹㄹ
-헐 설마 소고기넣고 끓인 라면...?
-탁무이 친구 연예인이었던거 몰랐던 사람 많았구나 ...ㄹㄹ 고딩때 탁무이 회상장면에서도 같이 촬영하고 그러는거 나왔었는데
-28화에 의사친구 연예인였다는거 나와여영수 다홍 의사친구랑 다 같이 아역으로 출연



수고 하셨습니다