

데이터 크롤링과 정제

동적 웹페이지 크롤링

목차

- 동적 웹페이지 크롤링 준비
 - selenium 라이브러리 설치 및 WebDriver다운로드
- 동적 웹페이지 크롤링
 - 커피빈 홈페이지
 - 매장 검색
 - 교보문고 베스트 셀러

동적 웹 페이지 크롤링 준비 #1

- Selenium 라이브러리 설치
 - > `conda install selenium`

```
changsu — python — bash — 85x27
(base) Changsuui-MacBook-Pro:~ changsu$ conda install selenium
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /Users/changsu/opt/anaconda3

added / updated specs:
- selenium

The following packages will be downloaded:



| package          | build             |        |
|------------------|-------------------|--------|
| selenium-3.141.0 | py39h9ed2024_1000 | 818 KB |
| Total:           |                   | 818 KB |



The following NEW packages will be INSTALLED:

selenium          pkgs/main/osx-64::selenium-3.141.0-py39h9ed2024_1000

Proceed ([y]/n)? y
```

동적 웹 페이지 크롤링 준비 #2

- selenium 설치 확인
 - `conda list selenium`

```
changsui — -bash — 85x27

The following packages will be downloaded:

package | build
-----|-----
selenium-3.141.0 | py39h9ed2024_1000 818 KB
-----|-----
Total: 818 KB

The following NEW packages will be INSTALLED:

selenium pkgs/main/osx-64::selenium-3.141.0-py39h9ed2024_1000

Proceed ([y]/n)? y

Downloading and Extracting Packages
selenium-3.141.0 | 818 KB | ##### | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
[(base) Changsuui-MacBook-Pro:~ changsu$ conda list selenium
# packages in environment at /Users/changsui/opt/anaconda3:
#
# Name Version Build Channel
selenium 3.141.0 py39h9ed2024_1000
(base) Changsuui-MacBook-Pro:~ changsu$
```

동적 웹 페이지 크롤링 준비 #3

■ 크롬 웹브라우저 버전 확인

- 크롬 웹브라우저 주소창에 <chrome://version> 입력

Chrome: 103.0.5060.114 (공식 빌드) (x86_64)  **chrome**

개정: alc2360c5b02a6d4d6ab33796ad8a268a6128226-refs/branch-heads/50600{#1124}

OS: macOS 버전 12.4(빌드 21F79)

JavaScript: V8 10.3.174.18

사용자 에이전트: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/103.0.0.0 Safari/537.36

명령줄: /Applications/Google Chrome.app/Contents/MacOS/Google Chrome --flag-switches-begin --flag-switches-end

실행 가능 경로: /Applications/Google Chrome.app/Contents/MacOS/Google Chrome

Google LLC
Copyright 2022 Google LLC. All rights reserved.

■ ChromeDriver 다운로드

- <https://chromedriver.chromium.org/downloads>
- 자신의 크롬 웹브라우저 버전에 맞는 ChromeDriver 다운로드

Current Releases

- If you are using Chrome version 104, please download [ChromeDriver 104.0.5112.29](#)
- If you are using Chrome version 103, please download [ChromeDriver 103.0.5060.53](#)
- If you are using Chrome version 102, please download [ChromeDriver 102.0.5005.61](#)
- For older version of Chrome, please see below for the version of ChromeDriver that supports it.

동적 웹 페이지 크롤링 준비 #4

- 사용하는 운영체제의 ChromeDriver 다운로드
 - 압축해제 후 특정 경로에 저장
 - Mac OS 사용자: `/usr/local/bin` 폴더에 저장
 - Windows 사용자: 임의의 폴더에 저장 후 경로를 지정해줌

```
from selenium import webdriver
```

```
driver = webdriver.Chrome('/usr/local/bin/chromedriver')
```

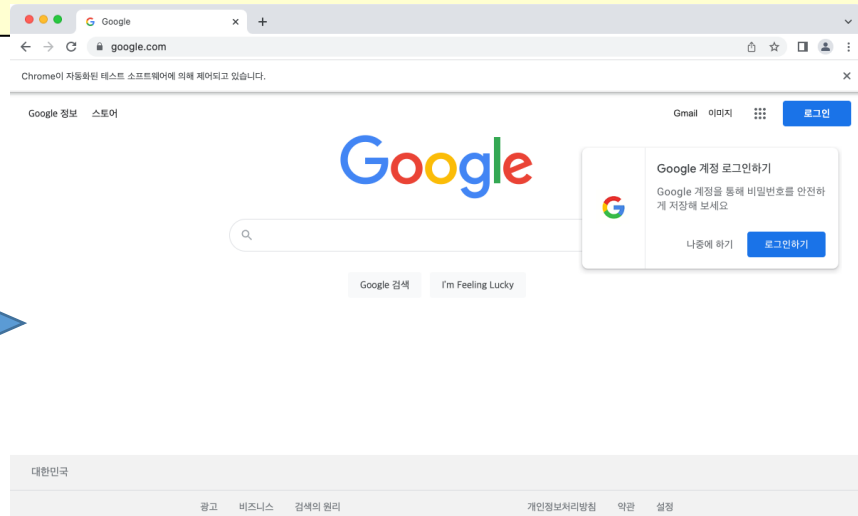
```
#driver = webdriver.Chrome('C:\\workspace') # Windows 사용자의 설치 경로
```

```
driver.get('https://www.google.com')
```

```
print(driver.current_url)
```

```
driver.quit() # driver종료
```

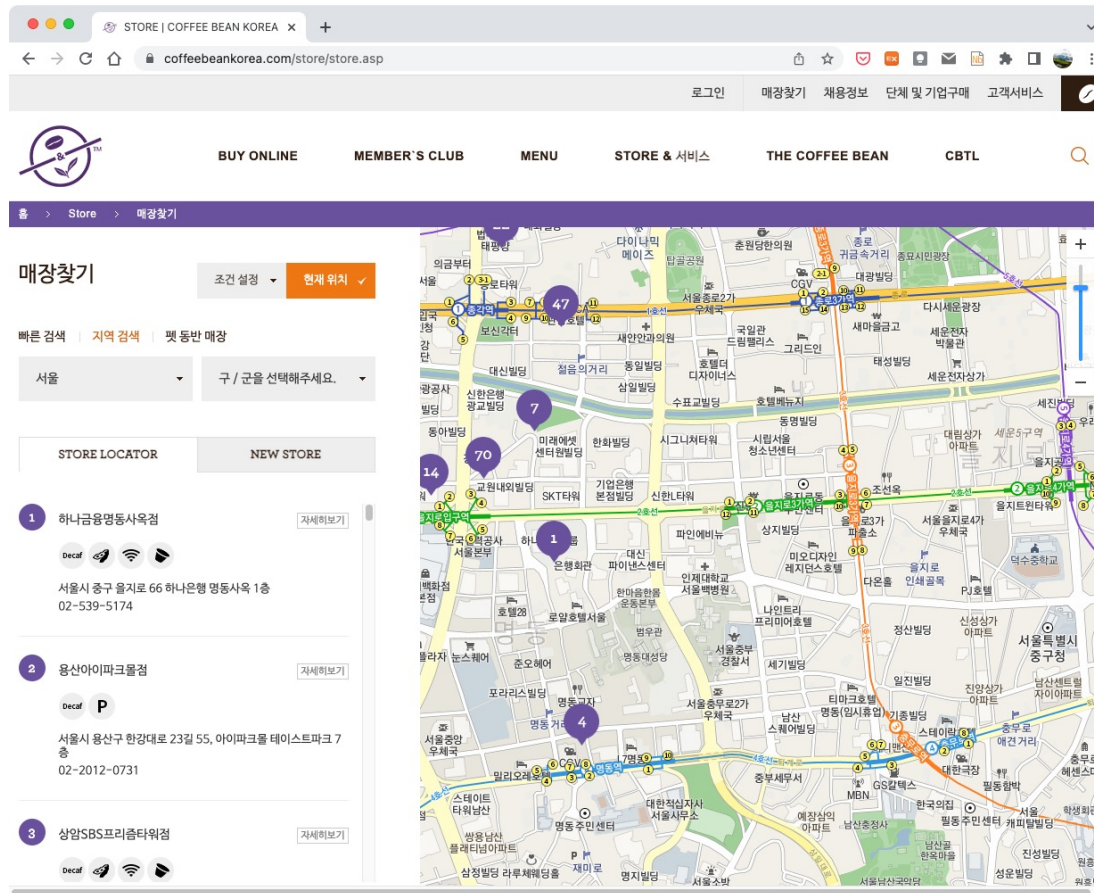
소스 코드를 실행하면
크롬 웹브라우저가
실행됨



동적 웹페이지 크롤링

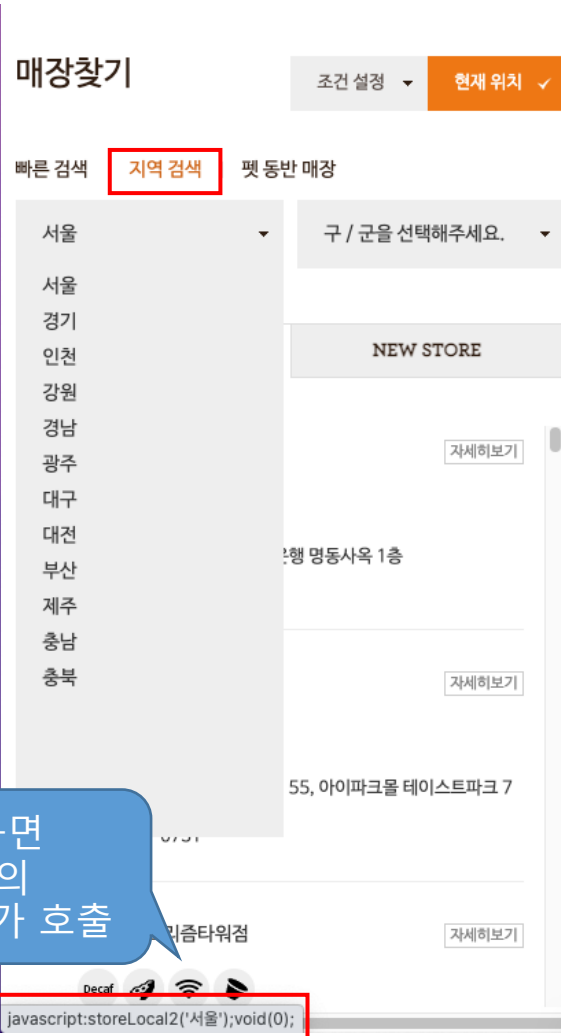
■ 웹페이지 분석하기

- Coffee Bean Korea 매장 찾기 url
 - <https://www.coffeebeankorea.com/store/store.asp>
- 지역 검색: 서울



동적 웹페이지 크롤링

- 매장 찾기 > 지역 검색 > '서울' 입력
 - javascript.storeLocal2('서울') 표시



시/도를 선택하면
자바스크립트의
storeLocal2() 함수가 호출



동적 웹페이지 크롤링

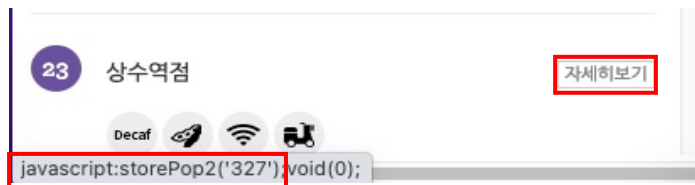
■ 소스 보기

- HTML 소스에는 조회된 매장 리스트가 없음
 - 시/도를 선택하면 storeLocal2()함수를 호출
 - 선택한 지역의 매장 목록을 표시하는 동적 웹 페이지 구조

```
<!-- store_tab -->
<div class="store_tab">
  <!-- STORE LOCATOR -->
  <h3 class="store_locator on"><a href="#" id="store_local">STORE LOCATOR</a></h3>
  <div class="cont">
    <ul id="storeListUL">
      <!-- 조회된 매장 리스트 -->
    </ul>
  </div>
</div>
```

■ 버튼에 연결된 자바스크립트 확인

- '자세히 보기' 버튼: javascript.storePop2('327')로 표시



상수역점

Detail

영업시간 평일 07:30~21:30 주말/공휴일 09:00~21:30

주최

주소 서울시 마포구 위암산로10길 3, 열케이빌딩 1~3층

전화번호 02-325-6804

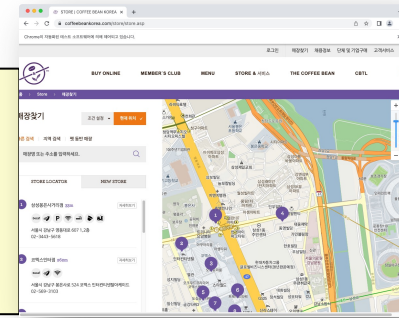
팝업창 표시

동적 웹페이지 크롤링 예제 코드 #1

■ 커피빈 코리아 홈페이지 자동 실행

```
from bs4 import BeautifulSoup
from selenium import webdriver

driver = webdriver.Chrome()
driver.get('https://www.coffeebeankorea.com/store/store.asp')
```

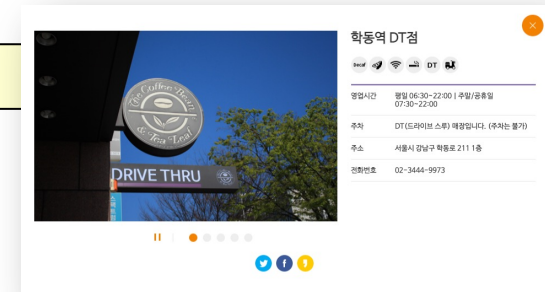


■ storePop2(1) 호출

- 팝업 창에 1번 매장인 '학동역 DT점' 나타남

```
driver.execute_script('storePop2(1)')
```

– 소스 보기에서도 매장정보가 보이지 않음



동적 웹페이지 크롤링 예제 코드 #2

■ 함수 호출 결과 페이지를 별도로 저장

```
html = driver.page_source # page_source: 해당 웹페이지의 소스가 저장됨
soup = BeautifulSoup(html, 'html.parser')

print(soup.prettify()) # HTML 소스를 보기 좋게 출력
```

• page_source

- requests.get() 함수를 사용해서 가져온 text 내용과 동일
- BeautifulSoup과 연동해서 필요한 정보를 크롤링

```
<div class="store_txt">
  <p class="name">
    <span>
      삼성봉은사거리점
      <strong class="distance">
        33m
      </strong>
    </span>
    <a class="btn_style6" href="javascript:storePop2('31');void(0);">
      자세히보기
    </a>
  </p>
  . . .
  <p class="address">
    <span>
      서울시 강남구 영동대로 607 1,2층
    </span>
  </p>
</div>
```

매장이름:
div.store_txt > p.name > span

매장주소:
p.address > span

예제 코드1: #1

- storePop2(1)을 실행 했을때 검색 결과 출력

```
from bs4 import BeautifulSoup
from selenium import webdriver
```

```
driver = webdriver.Chrome()
driver.get('https://www.coffeebeankorea.com/store/store.asp')
```

```
driver.execute_script('storePop2(1)')
```

execute_script(자바스크립트)

```
# 현재의 html 소스를 저장
html = driver.page_source
soup = BeautifulSoup(html, 'html.parser')
#print(soup.prettify()) # HTML 소스를 보기 좋게 출력
```

```
store_names = soup.select('div.store_txt > p.name > span')
store_name_list = []
for name in store_names:
    store_name_list.append(name.get_text())
```

```
print('매장 개수: ', len(store_name_list))
print(store_name_list) # 정상 동작
```

```
store_addresses = soup.select('p.address > span')
store_address_list = []
```

```
for addr in store_addresses:
    print(addr.get_text())
```

실행 결과

매장 개수: 34

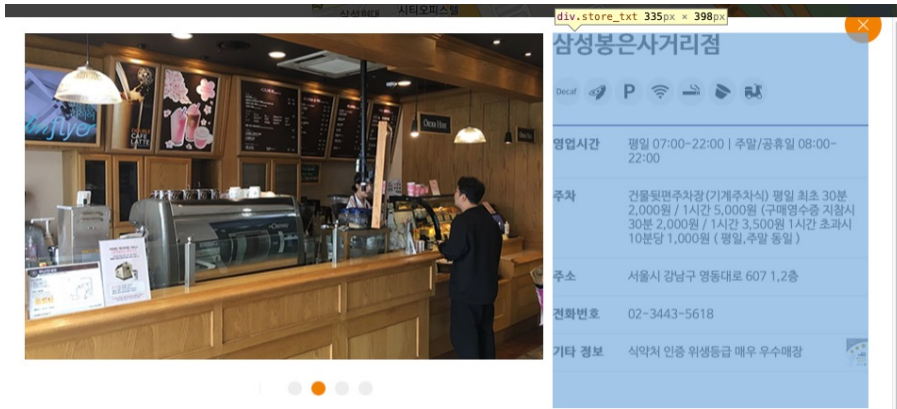
['삼성봉은사거리점33m', '코엑스인터점262m', '삼성오크우드점316m', '삼성봉은사로점442m', '도심공항타워점495m', '트레이드타워점507m', '도심공항점560m', 'CBTL(커피빈)파르나스몰점626m', '삼성로점632m', '삼성루첸타워점845m', '제주애월DT점', '하나금융명동사옥점', '진주NH본사점', '이천하이닉스 행복마을점', '용산아이파크몰점', '상암SBS프리즘타워점', '안양LS타워점', '명동거리점', '용인에버랜드점', '혜화역점', '제주아라점', '천안신부동점', '부산중앙역점', '용인역북점', '부산화명수정역점', '일산동구청점', '서초중앙로점', '청계천을지로점', '송파헬리오시티점', '수서역원방빌딩점', 'KB국민은행노원역점', '광주봉선점', '이태원몬드리안호텔점', '창원상남동점']

서울시 강남구 영동대로 607 1,2층
서울시 강남구 봉은사로 524 코엑스 인터컨티넨탈아케이드
서울 강남구 영동대로 513 컨벤션별관A동 지하 2층
서울시 강남구 봉은사로 628 엘스빌딩1층
서울시 강남구 테헤란로87길 36 공항타워 1층
서울시 강남구 영동대로 511 트레이드타워 지하1층
서울시 강남구 테헤란로87길22 도심공항터미널 지하1층
서울시 강남구 테헤란로521 파르나스몰 지하1층
서울시 강남구 삼성로 551 1층,2층
서울 강남구 대치동 943-2 루첸타워 1층 로비
제주도 제주시 애월읍 애월해안로 560
...

커피빈 지점 검색

■ 팝업 창 소스 확인

- `driver.execute_script('storePop2(%d)' % i)` 실행 시 팝업 창 HTML 소스



```
<div class="store_txt">
  <h2>삼성봉은사거리점</h2>
  <p class="tag">
    <span class="decaf" title="디카페인">디카페인</span>
    <span class="soy" title="두유">두유</span>
    <span class="parking" title="주차가능">주차가능</span>
    <span class="wifi" title="와이파이">와이파이</span>
    <span class="smoking" title="흡연가능">흡연가능</span>
    <span class="capsule" title="캡슐판매">캡슐판매</span>
    <span class="delivery" title="딜리버리">딜리버리</span>
  </p>
  <table class="store_table">
    <tbody>
      <tr>...</tr>
      <tr>...</tr>
      <tr>
        <th>주소</th>
        <td>...</td>
      </tr>
      <tr>
        <th>전화번호</th>
        <td>02-3443-5618</td>
      </tr>
      <tr>...</tr>
    </tbody>
  </table>
</div>
```

■ 지점 정보

- `<div class="store_txt">` 내부에 있음

- 지점명: `<h2>`

- 주소, 전화번호: `<table class="store_table">` 내부

커피빈 지점 검색 및 크롤링 정보를 저장하기

■ 자바스크립트 실행

```
for i in range(1, 380):  
    driver.get(coffeebean_url)  
    time.sleep(1) # 웹페이지를 연결할 동안 1초 대기  
  
    driver.execute_script('storePop2(%d)' % i) # 각 지점의 번호를 전달  
    time.sleep(1)
```

■ 지점 정보 가져오기

```
soup = BeautifulSoup(html, 'html.parser')  
# 매장 이름 검색  
store_name = soup.select_one('div.store_txt > h2').text  
print('매장이름: ', store_name)  
store_info = soup.select("div.store_txt > table.store_table > tbody > tr > td")  
  
store_address_list = list(store_info[2])  
  
store_addr = store_address_list[0] # 매장 주소  
store_phone = store_info[3].text # 매장 전화번호
```

동적 웹페이지 크롤링 예제 코드 #1

```
from bs4 import BeautifulSoup
from selenium import webdriver
import pandas as pd
import time

def coffeebean_store(store_list):
    coffeebean_url = 'https://www.coffeebeankorea.com/store/store.asp'
    driver = webdriver.Chrome()

    for i in range(1, 380):
        driver.get(coffeebean_url)
        time.sleep(1) # 웹페이지를 연결할 동안 1초 대기

        driver.execute_script('storePop2(%d)' % i)
        time.sleep(1)

        try:
            html = driver.page_source
            soup = BeautifulSoup(html, 'html.parser')

            store_name = soup.select_one('div.store_txt > h2').text # 매장 이름
            store_info = soup.select("div.store_txt > table.store_table > tbody > tr > td")
            store_address_list = list(store_info[2])
            store_addr = store_address_list[0] # 매장 주소
            store_phone = store_info[3].text # 매장 전화번호
            print('{} {} {}'.format(i+1, store_name, store_addr, store_phone))
            store_list.append([store_name, store_addr, store_phone])
        except:
            continue
```

execute_script(스크립트명):
javascript 실행

동적 웹페이지 크롤링 예제 코드 #4

```
def main():
    store_info = []
    coffeebean_store(store_info)
    # DataFrame으로 변경
    coffeebean_table = pd.DataFrame(store_info, columns=('매장이름', '주소', '전화번호'))
    print(coffeebean_table.head())

    # DataFrame을 csv파일로 저장 (utf-8로 인코딩)
    coffeebean_table.to_csv('coffeebean_branches.csv', encoding='utf-8', mode='w',
index=True)

main()
```

```
학동역 DT점 서울시 강남구 학동로 211 1층 02-3444-9973
차병원점 서울시 강남구 논현로 566 강남차병원1층 02-538-7615
강남대로점 서울시 서초구 강남대로 369 1층 02-588-5778
강남에스점 서울시 서초구 서초 1306-3호 02-593-5095
청담에스점 서울시 강남구 압구정로 461 네이처포엠빌딩B108,109호 02-548-6052
신사점 서울시 강남구 도산대로 126 02-548-2741
압구정역점 서울시 강남구 논현로 842 압구정빌딩1층 02-544-6823
...
```

교보문고 베스트셀러 크롤링

- 국내도서 > 컴퓨터/IT > 하위목록만 검색
 - <http://www.kyobobook.co.kr/categoryRenewal/categoryMain.1af?linkClass=3311&mallGb=KOR&orderClick=sgx>
 - linkClass=3311: 컴퓨터/IT > 데이터베이스
- 크롤링 항목

linkClass 번호	항목
3301	컴퓨터공학
3311	데이터베이스
3312	웹프로그래밍
3315	프로그래밍언어
3316	모바일프로그래밍

교보문고 베스트셀러 크롤링

- 각 항목별 다음 페이지 이동
 - 자바스크립트: `_go_targetPage(숫자)`

```
<div class="list_paging align_center" id="eventPaging">
<div class='list_paging'>
  <ul>
    <li>
      <strong><a href="javascript:_go_targetPage('1')">1</a></strong>
    </li>
    <li>
      <a href="javascript:_go_targetPage('2')">2</a>
    </li>
    . . .
    <li>
      <a href="javascript:_go_targetPage('7')">7</a>
    </li>
  </ul>
  <a class='btn_next' href="javascript:_go_targetPage('2')">
    
    </a>
  </div>
</div>
```

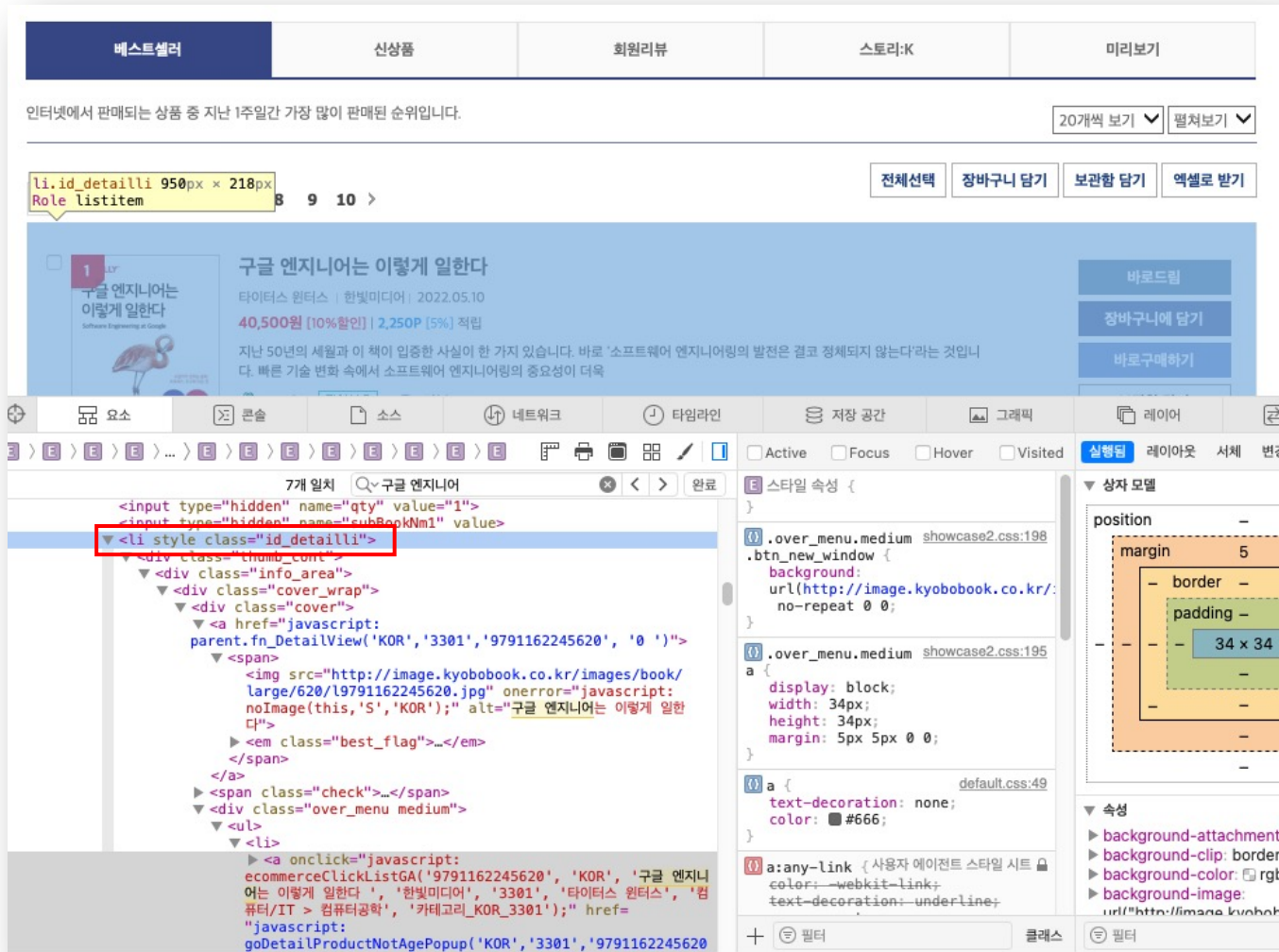
해당 카테고리의 전체
페이지 수 구분
- 의 개수로 구분 가능



교보문고 베스트 셀러 크롤링

■ 베스트 셀러 도서 정보

- `<li style class="id_detailli">` 내부에 존재함



교보문고 베스트 셀러 크롤링



구글 엔지니어는 이렇게 일한다

타이터스 원터스 | 한빛미디어 | 2022.05.10

40,500원 [10%할인] | **2,250P** [5%] 적립

지난 50년의 세월과 이 책이 입증한 사실이 한 가지 있습니다. 바로 '소프트웨어 엔지니어링의 발전은 결코 정체되지 않는다'라는 것입니다. 빠른 기술 변화 속에서 소프트웨어 엔지니어링의 중요성이 더욱

9.3 / 10 **잘읽혀요** 리뷰 20

[배송일정] 지금 주문하면 내일(12일, 화) 도착 예정

■ 도서 랭킹

```
<em class="best_flag"><span>1</span></em>
```

■ 책 제목

```
<div class="title">
  <a onclick="javascript:ecommerceClickListGA('9791162245620', 'KOR',
    '구글 엔지니어는 이렇게 일한다', '한빛미디어', '3301', '타이터스 원터스',
    '컴퓨터/IT &gt; 컴퓨터공학', '카테고리_KOR_3301');"
    href="javascript:makeDetailUrl('KOR','3301','9791162245620','0 ','N' )">
    <strong>구글 엔지니어는 이렇게 일한다</strong></a>
</div>
```

교보문고 베스트 셀러 크롤링



구글 엔지니어는 이렇게 일한다

타이터스 윈터스 | 한빛미디어 | 2022.05.10

40,500원 [10%할인] | 2,250P [5%] 적립

지난 50년의 세월과 이 책이 입증한 사실이 한 가지 있습니다. 바로 '소프트웨어 엔지니어링의 발전은 결코 정체되지 않는다'라는 것입니다. 빠른 기술 변화 속에서 소프트웨어 엔지니어링의 중요성이 더욱

9.3 / 10 **잘읽혀요** 리뷰 20

[배송일정] 지금 주문하면 내일(12일, 화) 도착 예정

■ 저자, 출판사, 출판일

```
<div class="pub_info">
  <span class="author">타이터스 윈터스</span>
  <span class="publication">한빛미디어</span>
  <span class="publication">
    2022.05.10 <!-- 초판일 -->
  </span>
</div>
```

■ 설명

```
<div class="info">
  <span>지난 50년의 세월과 이 책이 입증한 사실이 한 가지 있습니다. 바로
  '소프트웨어 엔지니어링의 발전은 결코 정체되지 않는다'라는 것입니다. 빠른 기술 변화
  속에서 소프트웨어 엔지니어링의 중요성이 더욱</span>
</div>
```

코드 설명

- itertools 모듈

- count(start, step) 함수
 - 기본값: start=0, step=1
 - start로부터 시작해서 step만큼 증가를 무한 반복

교보문고 베스트 셀러 1권 정보 분석

```
html = '''
<div class="title">
<a onclick="javascript:ecommerceClickListGA('9791162245620', 'KOR', '구글 엔지니어는 이렇게 일한다
', '한빛미디어', '3301', '타이터스 원터스', '컴퓨터/IT > 컴퓨터공학', '카테고리_KOR_3301');"
href="javascript:makeDetailUrl('KOR','3301','9791162245620','0 ','N ')">
<strong>구글 엔지니어는 이렇게 일한다</strong></a>
</div>
<div class="pub_info">
<span class="author">타이터스 원터스</span>
<span class="publication">한빛미디어</span>
<span class="publication">
2022.05.10 <!-- 초판일 --></span>
</div>
<div class="info">
<span>지난 50년의 세월과 이 책이 입증한 사실이 한 가지 있습니다. 바로 '소프트웨어 엔지니어링의 발전은
결코 정체되지 않는다'라는 것입니다. 빠른 기술 변화 속에서 소프트웨어 엔지니어링의 중요성이 더욱</span>
</div>
'''

from bs4 import BeautifulSoup
book = BeautifulSoup(html, 'html.parser')

book_title = book.find('div', attrs={'class': 'title'}).find('strong').text
book_info = book.find('div', attrs={'class': 'pub_info'})
book_author = book_info.find('span', attrs={'class': 'author'}).text
book_publication = book_info.find('span', attrs={'class': 'publication'}).text
book_summary = book.find('div', attrs={'class': 'info'}).find('span').text

print('title:', book_title)
print('author:', book_author)
print('publication:', book_publication)
print('summary:', book_summary)
```


교보문고 베스트 셀러 크롤링 소스 코드 #1

```
import time
from bs4 import BeautifulSoup
import pandas as pd
from itertools import count
from selenium import webdriver

...

category_num1 = [3301, 3302, 3303, 3305, 3307, 3309, 3310, 3311, 3312, 3313,
                 3314, 3315, 3316, 3317, 3319, 3321, 3323, 3325, 3328, 3329]

category_dic1 = {3301: '컴퓨터공학', 3302: 'IT일반', 3303: '컴퓨터입문_활용',
                 3305: '전산통계_해석', 3307: 'OS', 3309: '네트워크', 3310: '보안_해킹',
                 3311: '데이터베이스', 3312: '개발방법론', 3313: '게임',
                 3314: '웹프로그래밍', 3315: '프로그래밍언어', 3316: '모바일프로그래밍',
                 3317: 'OA_사무자동화', 3319: '웹사이트', 3321: '그래픽', 3323: '멀티미디어',
                 3325: 'CAD', 3328: '자격증_수험서', 3329: '대학교재'}

...

category_num = [3301, 3311, 3312, 3315, 3316]
category_dic = {3301: '컴퓨터공학', 3311: '데이터베이스', 3312: '웹프로그래밍', 3315:
'프로그래밍언어', 3316: '모바일프로그래밍'}
```

교보문고 베스트 셀러 크롤링 소스 코드 #2

```
def get_kobomungo_data():
    wd = webdriver.Chrome() # 본인의 webdriver 경로

    for category_idx in category_num:
        kobomungo_URL =
        "http://www.kyobobook.co.kr/categoryRenewal/categoryMain.laf?linkClass=%s&mallGb=KOR&orderCl
        ick=sgx" %str(category_idx)

        wd.get(kobomungo_URL)
        print("Category Index {}: {}".format(category_idx, category_dic.get(category_idx)))
        time.sleep(5)
        bestseller_list = []

        for page_idx in count():
            try:
                wd.execute_script("_go_targetPage('%s')" % str(page_idx + 1))
                print("Page index [%s]" % (str(page_idx + 1)))
            except Exception as e:
                break

        time.sleep(3)

        html = wd.page_source
        soupData = BeautifulSoup(html, 'html.parser')
        book_list = soupData.find_all('li', attrs={'class': 'id_detailli'})
```

execute_script(스크립트명):
javascript 실행

교보문고 베스트 셀러 크롤링 소스 코드 #3

```
for book in book_list:
    # <em class="best_flag"><span>1</span></em>
    book_rank = book.find('em', attrs={'class': 'best_flag'}).find('span').text
    book_title = book.find('div', attrs={'class': 'title'}).find('strong').text
    book_info = book.find('div', attrs={'class': 'pub_info'})
    book_author = book_info.find('span', attrs={'class': 'author'}).text
    book_publication = book_info.find('span', attrs={'class': 'publication'}).text
    book_summary = book.find('div', attrs={'class': 'info'}).find('span').text

    print(book_rank, book_title)
```

```
bestseller_list.append([book_rank, book_title, book_author,
                        book_publication, book_summary])
```

파일 저장

```
category_name = category_dic.get(category_idx) # category[0]: 분류번호
```

```
book_table = pd.DataFrame(bestseller_list,
                           columns=('순위', '제목', '저자', '출판사', '요약'))
```

```
book_table.to_csv("./%s분야_베스트셀러_list.csv" % str(category_name),
                  encoding="utf-8", mode='w', index=False)
```

```
def main():
    print('교보문고 베스트셀러 크롤링 시작 ')
    get_kobomungo_data()
    print('-----크롤링 종료 -----')
```

```
main()
```

교보문고 베스트 셀러 크롤링

■ 컴퓨터공학분야_베스트셀러_list.csv파일

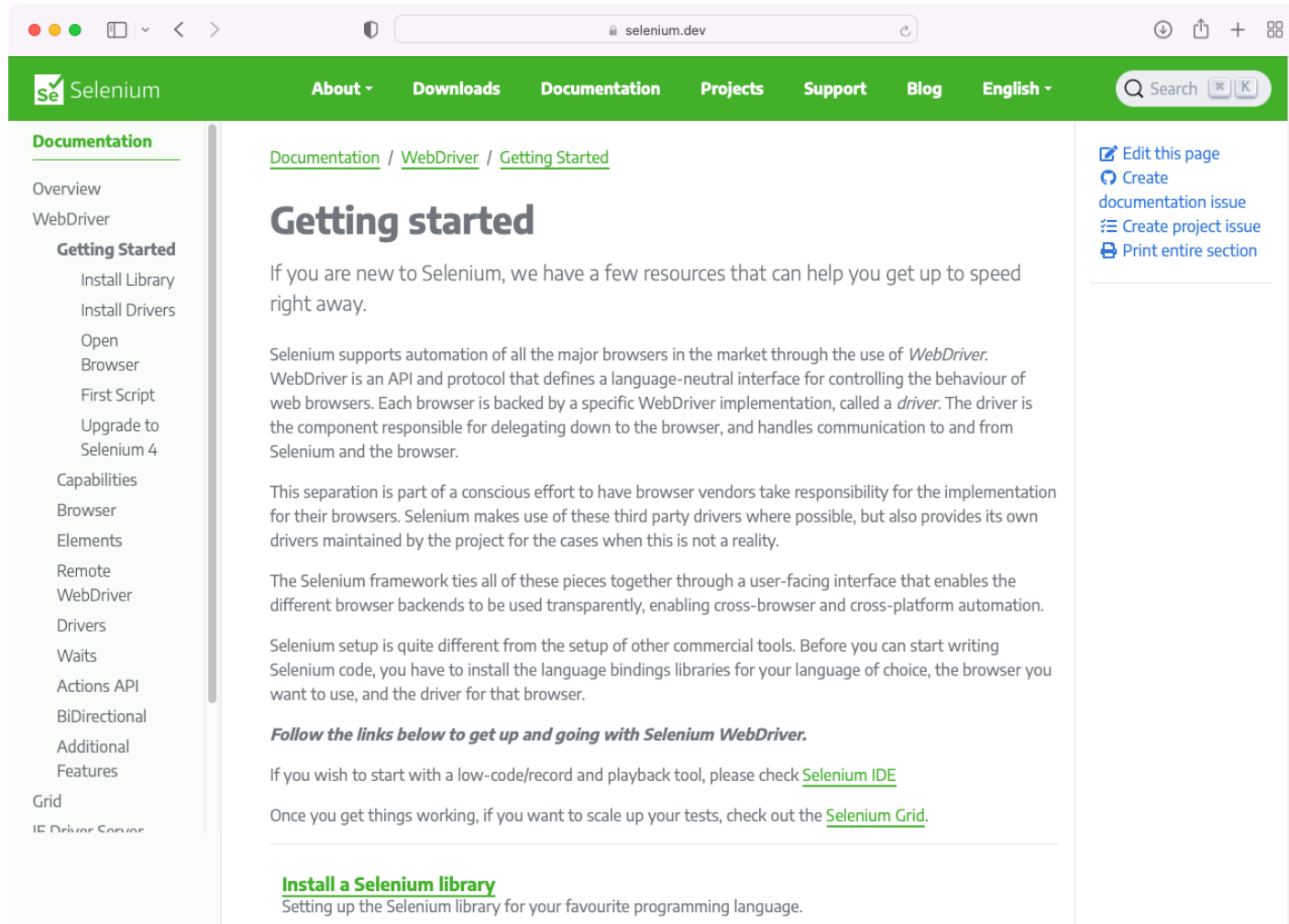
순위	제목	저자	출판사	요약
1	구글 엔지니어는 이렇게 일한다	타이터스 워터스	한빛미디어	지난 50년의 세월과 이 책이 입증한 사실이 한 가지 있습니다. 바로 '소프트웨어 엔지니어링'이
2	혼자 공부하는 머신러닝+딥러닝	박해선	한빛미디어	이 책은 수식과 이론으로 중무장한 머신러닝, 딥러닝 책에 지친 '독학하는 입문자'가 '꼭 필요
3	이펙티브 엔지니어	에드먼드 라우	길벗	뛰어난 엔지니어와 일반 엔지니어는 무엇이 다른가? 열심히 일하기와 똑똑하게 일하기는 어떻게 다른가? 구글, 페이스북, 인스타그램, 드롭박스 등 세계 최고 기업의 실제 사례 제시! 개인과 조직이 당장 &
4	밑바닥부터 시작하는 딥러닝	사이토 고키	한빛미디어	『밑바닥부터 시작하는 딥러닝』은 라이브러리나 프레임워크에 의존하지 않고, 딥러닝의 핵심을
5	클린 아키텍처: 소프트웨어 구조와 설계의 원칙	로버트 C. 마틴	인사이트	"살아있는 전설이 들려주는 실용적인 소프트웨어 아키텍처 원칙" 소프트웨어 아키텍처의 보편 원칙을 적용하면 소프트웨어 수명 전반에서 개발자 생산성을 획:
6	딥러닝 파이토치 교과서	서지영	길벗	머신 러닝 핵심 알고리즘부터 파이토치 기초, 합성곱 신경망, 설명 가능한 CNN, 순환 신경망
7	밑바닥부터 시작하는 딥러닝 2	사이토 고키	한빛미디어	이 책은 『밑바닥부터 시작하는 딥러닝』에서 다루지 못했던 순환 신경망(RNN)을 자연어 처리
8	핸즈온 머신러닝(2판)	오렐리아 제롱	한빛미디어	이 책은 지능형 시스템을 구축하려면 반드시 알아야 할 머신러닝, 딥러닝 분야 핵심 개념과 이
9	가상 면접 사례로 배우는 대규모 시스템 설계 기초(3판)	알렉스 쉬	인사이트	"페이스북의 뉴스 피드나 메신저, 유튜브, 구글 드라이브 같은 대규모 시스템은 어떻게 설계할까?" IT 경력자라도 느닷없이 대규모 시스템을 설계하려고 하면 막막하다고 느낄 수 있다.특히나 만
10	인공지능은 게임을 어떻게 움직이는가?	미야케 요이치로	이지스퍼블리싱	전략 게임에 사용하는 인공지능 기술 총망라! 플레이어를 몰입하게 만드는 전략 게임의 인공지능 기술을 파헤쳐보자! 국내에는 처음 소개되는 전략 게임의 인공지능 기술을 정리한 도서로, 전략 게임을 만드는
11	밑바닥부터 시작하는 딥러닝 3	사이토 고키	한빛미디어	이 책은 '밑바닥부터' 직접 만들어보며 즐겁게 딥러닝을 익히는 시리즈의 장점을 그대로 따왔
12	파이토치 딥러닝 마스터	엘리 스티븐스	책만	파이토치 핵심 개발진이 직접 집필한 책으로, 최고 실무자가 바로 옆에 앉아서 딥러닝의 기초! 텐서, 신경망 같은
13	Apache Airflow 기반의 데이터 파이프라인(I♥Clo	바스 하렌슬라	제이펍	이 책은 효과적인 데이터 파이프라인을 만들고 유지하는 방법을 설명하고 있으며, 이를 통해 (
14	데싸노트의 실전에서 통하는 머신러닝(Must Have)	권시현	골든레빗(주)	★ 실전은 피쳐 엔지니어링이다 ★ 뉴욕의 데이터 사이언티스트가 알려주는 ★ 머신러닝 문제풀이 해법을 익혀라 이 책은 실무와 캐글 챌린지에도 통하는 10가지 각 알고리즘을 엄선했다 머신러닝을 알려주
15	미래를 바꾼 아홉 가지 알고리즘(acornLoft)	존 맥코믹	에이콘출판	컴퓨터 세상을 만든 기발한 아이디어들 『미래를 바꾼 아홉 가지 알고리즘』. 오늘날 우리는 공
16	알고리즘 문제 해결 전략 세트(프로그래밍 대회에서	구중만	인사이트	프로그래밍 대회에서 배우는『알고리즘 문제 해결 전략 세트』, 프로그래밍 대회 문제를 풀면서
17	한 권으로 읽는 컴퓨터 구조와 프로그래밍	조너선 스타인하트	책만	컴퓨터는 어떻게 동작하며, 컴퓨터 프로그래밍이란 무엇인가? 개발자가 되기 위해 프로그래
18	모두의 딥러닝(3판)	조태호	길벗	비전공자 출신으로 오랜 기간 딥러닝 연구를 해온 저자는 자신이 겪어온 오류들을 겪지 않고,

참고자료

참고 자료

- Selenium documentation

- <https://www.selenium.dev/documentation/>





Questions?