

230303 파이썬 웹 크롤링

스타벅스 크롤링 & 분석

```
# 스타벅스 크롤링 분석
from selenium import webdriver
import time
import requests
from bs4 import BeautifulSoup as bs
from selenium.webdriver.common.by import By

url = 'https://www.starbucks.co.kr/store/store_map.do'
driver = webdriver.Chrome('chromedriver.exe')
driver.get(url)

area_btn = '#container > div > form > fieldset > div > section > article.find_store_co
nt > article > header.loca_search > h3 > a'
driver.find_element(By.CSS_SELECTOR, area_btn).click()
time.sleep(2)

seoul_btn = '#container > div > form > fieldset > div > section > article.find_store_c
ont > article > article:nth-child(4) > div.loca_step1 > div.loca_step1_cont > ul > li:
nth-child(1) > a'
driver.find_element(By.CSS_SELECTOR, seoul_btn).click()
time.sleep(2)

all_btn = '#mCSB_2_container > ul > li:nth-child(1) > a'
driver.find_element(By.CSS_SELECTOR, all_btn).click()
time.sleep(2)

txt = driver.page_source
html = bs(txt)

# 서울은 총 600개인데 내 주역 주변 스타벅스 까지 같이 나옴
len(html.select('li.quickResultLstCon')) # 610개

shops = html.select('li.quickResultLstCon')
shop = shops[0]
shop

# 위도
lat = shop['data-lat']
print(lat)
# 경도
lng = shop['data-long']
print(lng)
# 지점명
name = shop['data-name']
```

```

print(name)
# 주소
addr = str(shop.select('p.result_details')[0]).split('<br/>')[0].split('>')[1]
print(addr)
# 전화번호
tel = str(shop.select('p.result_details')[0]).split('<br/>')[1].split('<')[0]
print(tel)
# 매장종류
shoptype = shop.select('i.pin_generalDT')[0].text
print(shoptype)

```

```

# 기존의 데이터프레임에서 열을 하나 추가
df['시군구명'] = gu_names
df

```

	매장명	위도	경도	주소	전화번호	매장종류	시군구명
0	역삼아레나빌딩	37.501087	127.043069	서울특별시 강남구 연주로 425 (역삼동)	1522-3232	리저브 매장 2번	강남구
1	논현역사거리	37.510178	127.022223	서울특별시 강남구 강남대로 538 (논현동)	1522-3232	리저브 매장 2번	강남구
2	신사역성일빌딩	37.513931	127.020606	서울특별시 강남구 강남대로 584 (논현동)	1522-3232	리저브 매장 2번	강남구
3	국기원사거리	37.499517	127.031495	서울특별시 강남구 테헤란로 125 (역삼동)	1522-3232	리저브 매장 2번	강남구
4	대치재경빌딩R	37.494668	127.062583	서울특별시 강남구 남부순환로 2947 (대치동)	1522-3232	리저브 매장 2번	강남구
...
595	사가정역	37.579594	127.087966	서울특별시 중랑구 면목로 310	1522-3232	리저브 매장 2번	중랑구
596	상봉역	37.596890	127.086470	서울특별시 중랑구 망우로 307 (상봉동)	1522-3232	리저브 매장 2번	중랑구
597	목동	37.615368	127.076633	서울특별시 중랑구 동일로 952 (목동, 로프트원 테릉입구역) 1층	1522-3232	리저브 매장 2번	중랑구
598	양원역	37.606654	127.106360	서울특별시 중랑구 양원역로10길 3 (망우동)	1522-3232	리저브 매장 2번	중랑구
599	중화역	37.601709	127.078411	서울특별시 중랑구 봉화산로 35	1522-3232	리저브 매장 2번	중랑구

```

# '구'별 스타벅스 매장 수 계산

starbucks_count = df.pivot_table(index='시군구명', values='매장명', aggfunc='count')
starbucks_count

✓ 0.0s

```

매장명	
시군구명	
강남구	87
강동구	17
강북구	6
강서구	25
관악구	12
광진구	19
구로구	14
금천구	13
노원구	14
도봉구	4
동대문구	10
동작구	11
마포구	35
서대문구	22
서초구	48
성동구	15
성북구	15
송파구	34
양천구	17
영등포구	41
용산구	25
은평구	13
종로구	40
중구	55
중랑구	8

깃허브 python 230303 참고 !!