

Selenium 활용 데이터 수집 - 요일별 네이버 웹툰 목록

개요

selenium

웹 브라우저를 제어할 수 있는 파이썬 패키지

chromedriver

selenium과 Google Chrome 브라우저를 연결하는 프로그램

파이썬 스스로 chromedriver를 내려받도록 하기 위해 `chromedriver_autoinstaller` 패키지가 필요함

설치해야 하는 패키지

```
$ pip install --upgrade chromedriver_autoinstaller
$ pip install --upgrade selenium
```

#01. 패키지 참조

```
# ChromeDriver 자동 설치 모듈
import chromedriver_autoinstaller
# Chrome을 제어하기 위한 객체
from selenium import webdriver
# Chrome이 웹 페이지 로딩을 완료 할 때까지 최대 n초간 대기하는 기능.
from selenium.webdriver.support.ui import WebDriverWait
from bs4 import BeautifulSoup
from pandas import DataFrame
# 파이썬 프로그램에 지정된 시간동안 랙을 거는 기능을 위해 사용
import time
```

#02. 크롬브라우저 가동하기

```
# 크롬드라이버 자동 설치
chromedriver_autoinstaller.install()

# 크롬드라이버를 통해 크롬을 실행시킴
# -> driver 객체는 Chrome 자체
driver = webdriver.Chrome()

# 크롬브라우저가 준비될 때 까지 최대 5초씩 대기
driver.implicitly_wait(5)
```

```
# 요일별 네이버 웹툰에 접근하기 위한 변수값
# -> URL을 분석하여 얻어낸 값
```

```
params = ['mon', 'tue', 'wed', 'thu', 'fri', 'sat', 'sun', 'dailyPlus']
```

```
# 네이버 웹툰의 주소 형식
```

```
naverWebtoonUrl = "https://comic.naver.com/webtoon?tab={0}"
```

```
# 수집된 결과가 누적될 빈 리스트
```

```
naverWebtoonData = []
```

```
# 요일별 반복
```

```
for p in params:
```

```
    # 특정 요일의 네이버 웹툰 페이지
```

```
    url = naverWebtoonUrl.format(p)
```

```
    print(url)
```

```
    # 크롬브라우저로 방문시킴
```

```
    driver.get(url)
```

```
    time.sleep(1)
```

```
    # 브라우저에 표시되는 전체 코드를 추출 --> bs4 객체로 변환
```

```
    soup = BeautifulSoup(driver.page_source)
```

```
    #print(soup)
```

```
    # 웹툰이 표시되는 부분만 추출
```

```
    webtoonList = soup.select(".ContentList__content_list--q5KXY > .item")
```

```
    #print(webtoonList)
```

```
    # 추출된 웹툰 목록 수 만큼 반복
```

```
    for w in webtoonList:
```

```
        # 포스터 URL 가져오기
```

```
        poster = w.select(".Poster__image--d9XTI")
```

```
        # 가져온 이미지가 존재하고, src속성이 있다면?
```

```
        if poster and "src" in poster[0].attrs:
```

```
            posterValue = poster[0].attrs['src']
```

```
        else:
```

```
            posterValue = None
```

```
        #print(posterValue)
```

```
        # 웹툰의 URL 가져오기
```

```
        url = w.select(".Poster__link--sopnC")
```

```
        if url and "href" in url[0].attrs:
```

```
            urlValue = url[0].attrs['href']
```

```
            if urlValue.find("https://comic.naver.com") == -1:
```

```
                urlValue = "https://comic.naver.com" + urlValue
```

```
        else:
```

```
            urlValue = None
```

```
        #print(urlValue)
```

```
        # 웹툰 제목 가져오기
```

```
        title = w.select(".ContentTitle__title--e3qXt > .text")
```

```
        if title:
```

```

        titleValue = title[0].text.strip()
    else:
        titleValue: None
    #print(titleValue)

    # 작가 이름 가져오기
    author = w.select(".ContentAuthor__author--CTAAP")

    if author:
        authorValue = author[0].text.strip()
    else:
        authorValue = None
    #print(authorValue)

    # 별점 가져오기
    rating = w.select(".Rating__star_area--dFzsb > .text")

    if rating:
        ratingValue = rating[0].text.strip()
    else:
        ratingValue = None
    #print(ratingValue)

    # 결과 병합
    resultDic = {"요일": p, "제목": titleValue, "작가": authorValue, "별점": ratingV
    #print(resultDic)
    naverWebtoolData.append(resultDic)

df = DataFrame(naverWebtoolData)
df.to_excel("요일별_네이버_웹툰.xlsx")
df

```

<https://comic.naver.com/webtoon?tab=mon>

9.81
 9.88
 9.76
 9.98
 9.84
 9.83
 9.93
 9.92
 9.95
 9.91
 9.97
 9.96
 9.58
 9.94
 9.98
 9.52
 9.96
 9.96
 9.94
 9.97
 9.96
 9.93

9.92
9.96
9.96
9.95
9.97
9.88
9.97
9.86
9.92
9.98
9.97
9.97
9.93
9.94
9.91
6.23
9.87
8.94
9.91
9.85
9.95
9.16
9.81
9.97
9.88
9.93
9.94
9.93
9.80
9.87
9.95
9.94
9.96
9.94
9.96
9.96
9.95
9.80
9.88
9.91
9.94
9.96
9.92
9.93
9.97
9.91
9.93
9.96
9.89
9.95
9.82
9.93
9.96
9.85
9.92
9.95

9.76
9.95
9.91
9.96
9.93
9.85
9.85
9.89
9.02
9.95
9.85
<https://comic.naver.com/webtoon?tab=tue>
9.45
9.95
9.98
9.97
3.55
9.91
9.91
9.76
9.97
9.90
9.45
9.83
9.83
9.73
9.97
9.96
9.90
9.95
9.98
9.98
9.98
9.86
9.94
9.97
9.91
9.24
9.90
9.83
9.95
9.92
9.97
9.97
9.97
9.97
8.69
9.90
9.93
9.74
9.97
9.94
9.97
9.91
9.87
9.94

9.80
9.96
9.95
9.68
8.67
9.97
9.82
9.73
9.97
9.97
9.97
9.84
9.95
9.87
9.94
9.94
9.92
9.81
9.97
9.91
9.80
9.97
9.96
9.85
9.86
8.81
9.86
9.83
9.75
9.94
9.77
9.85
9.83

<https://comic.naver.com/webtoon?tab=wed>

9.92
9.94
9.94
9.67
8.70
9.82
9.95
9.94
9.85
9.97
9.97
9.79
9.30
9.92
9.98
9.91
9.96
9.66
9.86
9.28
9.98
9.91

9.97	
9.91	
8.52	
9.92	
9.97	
9.78	
9.97	
9.96	
9.95	
9.76	
9.95	
9.96	
9.84	
9.98	
9.96	
9.92	
9.91	
9.96	
9.69	
9.72	
9.96	
9.86	
9.97	
9.94	
8.69	
9.91	
9.96	
9.92	
9.97	
9.91	
9.97	
9.64	
9.89	
9.97	
9.88	
9.94	
9.96	
9.94	
9.94	
9.73	
9.95	
9.46	
9.95	
6.78	
9.93	
9.93	
9.69	
9.96	
9.91	
9.91	
9.80	
9.85	
9.88	
9.69	
9.92	
9.92	

9.90

9.86

9.83

9.89

9.60

9.84

9.92

9.87

9.74

9.04

9.86

9.63

9.37

<https://comic.naver.com/webtoon?tab=thu>

9.92

9.91

9.85

9.97

9.96

9.93

9.81

9.56

9.97

9.94

9.96

9.93

9.98

9.95

9.12

9.93

9.98

9.93

9.92

9.86

9.96

9.85

9.96

9.97

9.59

9.92

9.97

9.97

9.93

9.91

9.97

9.87

9.97

9.94

9.85

9.94

9.97

9.90

9.87

9.92

6.84

9.91

9.96
9.96
9.84
9.94
9.97
9.97
9.92
9.69
9.86
9.88
9.73
9.97
9.91
9.88
9.70
9.82
9.84
9.95
9.81
9.91
9.71
9.95
4.15
9.97
9.79
9.93
9.97
9.89
8.95
9.90
9.81
9.85
8.80
9.66
8.35
9.94
9.51
9.94
9.81
9.82
9.51
9.95
9.73
9.94
<https://comic.naver.com/webtoon?tab=fri>
9.47
9.91
9.97
9.98
9.96
9.97
9.97
9.96
9.94
9.98
9.97

9.90
9.96
9.98
9.77
9.95
9.95
9.97
9.90
9.97
9.87
9.97
9.79
9.93
9.77
9.93
9.97
9.96
9.48
9.92
9.97
9.92
9.96
9.97
9.86
9.97
9.98
9.97
9.84
9.94
9.91
9.48
8.43
9.97
9.76
9.96
9.86
9.97
9.97
9.90
9.94
9.87
8.90
9.82
9.96
9.97
9.81
9.95
9.90
9.95
9.96
9.97
9.92
9.95
9.96
9.63
9.85

9.85
9.78
9.84
9.93
9.84
9.96
9.83
9.93
9.96
9.93
9.94
9.73
9.91
9.90
9.71
9.73
9.95
9.63
9.65
9.92
9.91
9.89
9.28
<https://comic.naver.com/webtoon?tab=sat>
9.92
9.81
9.87
9.98
9.79
9.83
9.85
9.96
9.90
9.88
9.96
9.96
9.95
9.96
9.91
9.95
9.96
9.88
9.96
9.96
9.93
9.92
9.85
9.91
9.94
9.97
9.70
9.68
9.85
9.93
9.64
9.97

9.71
9.44
9.97
9.92
9.97
9.89
9.82
9.98
9.19
9.97
9.86
9.98
9.83
9.90
9.76
9.97
9.98
9.69
7.72
9.51
9.86
9.76
9.97
9.57
9.84
9.97
9.94
9.97
9.93
9.94
9.94
9.95
9.91
9.95
9.95
9.76
9.97
9.97
9.94
9.97
9.94
9.91
9.96
9.84
9.96
8.45
9.93
<https://comic.naver.com/webtoon?tab=sun>
9.67
9.72
9.76
9.89
9.94
9.68
9.56
9.98

9.96
9.91
9.95
9.15
9.92
9.89
9.94
9.95
9.71
9.98
9.97
9.97
9.93
9.86
9.88
9.98
9.83
9.98
9.90
9.86
9.96
9.97
9.94
9.94
9.98
9.92
9.89
9.97
9.97
9.88
9.77
9.91
9.79
9.92
9.93
9.95
9.97
9.92
9.97
9.97
9.91
9.95
9.96
9.93
9.86
9.95
9.96
9.97
9.96
9.97
9.92
9.91
9.55
9.73
9.96
9.96

9.96
9.96
9.96
9.97
9.75
9.53
9.96
9.89
9.91
9.86
9.80
9.86
9.92
9.86
9.90
9.81
9.95
9.72
9.83
9.58
9.89
9.39
9.77
<https://comic.naver.com/webtoon?tab=dailyPlus>
9.78
9.23
9.91
9.92
9.92
9.97
9.89
9.82
9.97
9.88
9.47
9.68
9.93
9.84
9.79
9.90
9.93
9.71
9.84
9.96
9.91
9.91
8.78
9.92
9.82
9.95
9.50
9.86
9.91
9.49
9.95
9.95

9.70	
9.77	
9.62	
9.88	
9.28	
9.93	
9.88	
9.80	
9.80	
9.75	
9.88	
9.83	
9.85	
9.95	
9.95	
5.26	
9.85	
9.94	
9.19	
9.93	
9.94	
9.96	
9.85	
9.88	
8.47	
9.85	
9.53	
9.60	
9.65	
9.88	
9.82	
9.76	
9.92	
9.93	
9.95	
9.89	
9.72	
9.90	
9.49	
9.94	
9.81	
9.67	
9.95	
9.33	
9.94	
9.96	
9.91	
8.89	
9.89	
9.80	
9.40	
9.92	
9.92	
9.88	
9.80	
9.68	

9.74
 9.94
 9.76
 9.95
 9.94
 9.92
 9.64
 9.95
 9.97
 9.66
 9.09
 9.07
 9.84
 9.52
 9.65
 9.58
 8.96
 9.79
 9.93
 9.77
 9.81
 9.68
 9.82
 9.98
 9.90
 9.58
 9.66
 9.83
 9.88
 9.83
 9.63
 9.37
 9.85
 8.83
 8.27
 9.53
 9.69
 9.33
 9.76

	요일	제목	작가	별점	포스터	
0	mon	뷰티풀 군바리	설이 / 윤성원	9.81	https://image-comic.pstatic.net/webtoon/648419...	https://comic.naver.com/webtoon/detail.naver?cid=648419&titleId=6...
1	mon	신의 탑	SIU	9.88	https://image-comic.pstatic.net/webtoon/183559...	https://comic.naver.com/webtoon/detail.naver?cid=183559&titleId=1...
2	mon	퀘스트	박태준 만화회사	9.76	https://image-comic.pstatic.net/webtoon/783052...	https://comic.naver.com/webtoon/detail.naver?cid=783052&titleId=7...

	요일	제목	작가	별점	포스터	
		지상주의				
3	mon	똑 님은 딸	이담	9.98	https://image-comic.pstatic.net/webtoon/774866...	https://comic.naver.com/webtoon/titleId=7...
4	mon	백수 세끼	치즈	9.84	https://image-comic.pstatic.net/webtoon/733074...	https://comic.naver.com/webtoon/titleId=7...
...
721	dailyPlus	후드	후렛샤 / 심규호	8.27	None	https://comic.naver.com/webtoon/titleId=8...
722	dailyPlus	실버벨	노떼	9.53	None	https://comic.naver.com/webtoon/titleId=7...
723	dailyPlus	씨앗의 정원	백희	9.69	None	https://comic.naver.com/webtoon/titleId=8...
724	dailyPlus	도와줘 우주	설레임 / 한겨울	9.33	None	https://comic.naver.com/webtoon/titleId=7...
725	dailyPlus	투두리스트	RYO / FLEWing	9.76	None	https://comic.naver.com/webtoon/titleId=7...

726 rows × 6 columns