

Crawling Day 2

Genie

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

#1~50
driver = webdriver.Chrome('chromedriver.exe')
url = 'https://www.genie.co.kr/chart/top200'
driver.get(url)
txt = driver.page_source
html = bs(txt)

songs = html.select('tbody>tr')

song_data = []
rank = 1
for song in songs:
    title = song.select('a.title')[0].text.strip()
    singer = song.select('a.artist')[0].text
    song_data.append(['Genie', rank, title, singer])
    rank += 1

#51~100
url = 'https://www.genie.co.kr/chart/top200?ditc=D&ynd=20230227&hh=16&rtm=Y&pg=2'
driver.get(url)
txt = driver.page_source
html = bs(txt)

songs = html.select('tbody>tr')

for song in songs:
    title = song.select('a.title')[0].text.strip()
    singer = song.select('a.artist')[0].text
    song_data.append(['Genie', rank, title, singer])
    rank += 1
df = pd.DataFrame(song_data, columns=['서비스', '순위', '타이틀', '가수'])
df
df.to_excel('genie.xlsx', index=False)
```

<https://s3-us-west-2.amazonaws.com/secure.notion-static.com/a7601187-0529-4bf6-9080-3515d64c6002/genie.xlsx>

