

```
1 Exercise. Python Regular Expression1
2
3 import requests
4 from bs4 import BeautifulSoup
5 # 위의 두 모듈이 없는 경우에는 pip install requests bs4 실행
6
7 def get_news_content(url):
8     response = requests.get(url)
9     content = response.text
10
11     soup = BeautifulSoup(content, 'html5lib')
12
13     div = soup.find('div', attrs = {'id' : 'harmonyContainer'})
14     content = ""
15     for paragraph in div.find_all('p'):
16         content += paragraph.get_text()
17
18     return content
19
20 news1 = get_news_content('https://news.v.daum.net/v/20190617073049838')
21 print(news1)
22
23
24 #Microsoft Visual Code 혹은 일반 editor로 실행시 아래와 같은 error가 발생할 수 있다.
25 #requests.exceptions.SSLError: HTTPSConnectionPool(host='news.v.daum.net', port=443): Max retries exceeded with url:
26 #/v/20190617073049838 (Caused by SSLError("Can't connect to HTTPS URL because the SSL module is not available."))
27 #해결방법
28 #아래의 사이트로 가서 'Win64OpenSSL-1_1_1c.exe'를 다운로드하여 설치한다.
29 #https://slproweb.com/download/Win64OpenSSL-1\_1\_1c.exe
```