```
# 받은 응답(HTML 코드)를 파싱하기 위해 'BeautifulSoup'을 임포트
from bs4 import BeautifulSoup
# 타겟 URL에 요청을 보내고 응답을 받기 위해 'urllib'
import urllib.request
# quote'는 'urlopen'에서 인자로 사용되는 URL주소(이하 타겟 주소)에 한글(UTF_8)
# 이 포함되었을 때,# 이를 아스키(ASCII)형식으로 바꿔주기 위한 함수
from urllib.parse import quote
keyword = input('검색어:')
page_num = int(input('페이지 개수(페이지 당 15개):'))
output_file_name = input("저장할 파일명:")
output_file = open(output_file_name, 'w', encoding="utf-8")
# 'check_news=1'은 뉴스로 한정 지은 것을, 'more=1'은 더 보기를 누른 상태임을,
# 'sorting=3'과 'range=3'은 각각 정확도 순 정렬이며, 전체기간검색임
for i in range(page_num):
    current_page_num = 1 + i * 15
    target_URL = "http://news.donga.com/search?p=" + str(current_page_num) + \
          '&query=' +quote(keyword) + \
          '&check_news=1&more=1&sorting=3&search_date=1&range=3'
    source_code_from_URL = urllib.request.urlopen( target_URL)
    soup = BeautifulSoup(source_code_from_URL, 'lxml', from_encoding='utf-8')
    for title in soup.find_all('p', 'tit'):
        title_link = title.select('a')
        article_URL = title_link[0]['href']
        source_code_from_url = urllib.request.urlopen(article_URL)
        soup = BeautifulSoup(source_code_from_url, 'lxml',from_encoding='utf-8')
        content_of_article = soup.select('div.article_txt')
        for item in content_of_article:
            string_item = str(item.find_all(text=True))
            output_file.write(string_item)
output_file.close()

<Histogram>

import matplotlib.pyplot as plt
import pandas as pd
from matplotlib import font_manager, rc
font_name = \
    font_manager.FontProperties(fname="c:/Windows/Fonts/malgun.ttf").get_name()
rc('font', family=font_name)
df = pd.read_csv('csv_exam1.csv',encoding='ms949')
data = pd.concat([df['국어'],df['영어'],df['수학']])
plt.hist(data, bins=3)
plt.xticks(range(0,100,40),['하', '중', '상'])
plt.title('점수빈도')
plt.show()

import matplotlib.pyplot as plt
import pandas as pd
from matplotlib import font_manager, rc
font_name = \
    font_manager.FontProperties(fname="c:/Windows/Fonts/malgun.ttf").get_name()
rc('font', family=font_name)
```

```
df = pd.read_csv('csv_exam1.csv',encoding='ms949')
plt.hist((df['국어'],df['영어'],df['수학']), bins=10, label=('국어','영어','수학'))
plt.title('점수빈도')
plt.legend()
plt.show()

<bar>
import matplotlib.pyplot as plt
import pandas as pd
from matplotlib import font_manager, rc
font_name = \
    font_manager.FontProperties(fname="c:/Windows/Fonts/malgun.ttf").get_name()
rc('font', family=font_name)
df = pd.read_csv('korea.csv',encoding='ms949')
print(df)
plt.figure()
plt.bar(df.index, df['점수'],width=1.0, color='r')
plt.xticks(range(0,len(df.index),1),df['이름'], rotation='vertical')
plt.title('학생별 국어 점수')
plt.show()

import matplotlib.pyplot as plt
import pandas as pd
from matplotlib import font_manager, rc
font_name = \
    font_manager.FontProperties(fname="c:/Windows/Fonts/malgun.ttf").get_name()
rc('font', family=font_name)
df = pd.read_csv('csv_exam1.csv',encoding='ms949')
print(df)
plt.figure()
plt.barh(df.index, df['국어'], color='r', label='국어')
plt.barh(df.index, -df['영어'], color='g', label='영어')
plt.title('학생별 국어,영어 점수')
plt.yticks(range(0,len(df.index),1),df['이름'], rotation='horizontal')
plt.xticks([-100,-50,0,50,100],(100,50,0,50,100))
plt.legend()
plt.show()

<line>
import matplotlib.pyplot as plt
from pandas import Series, DataFrame
s = Series([84900, 818000, 1756,292000])
# 객체생성
plt.figure()
# 출력
plt.plot(s)
plt.show()

import matplotlib.pyplot as plt
from pandas import Series
s1 = Series([84900, 81800, 71756, 92000]) #Series
s2 = Series([80500, 82000, 71736, 90000]) #Series
```

```
105 plt.figure(figsize=(10,4))
106 plt.plot(s1, label='04-10')
107 plt.plot(s2, label='04-11')
108 plt.grid()
109 plt.xlabel('index')
110 plt.ylabel('stock')
111 plt.title('plot graph')
112 plt.legend()
113 plt.show()
114
115 <box>
116 import matplotlib.pyplot as plt
117 import pandas as pd
118
119 from matplotlib import font_manager, rc
120 font_name = \
121     font_manager.FontProperties(fname="c:/Windows/Fonts/malgun.ttf").get_name()
122 rc('font', family=font_name)
123 df = pd.read_csv('csv_exam1.csv',encoding='ms949')
124 print(df)
125 plt.boxplot((df['국어'],df['영어'],df['수학']), labels=('국어','영어','수학'))
126 print(df['수학'].min())
127 print(df['수학'].mean())
128 print(df['수학'].median())
129 plt.title('점수분포')
130 plt.show()
131
132 <scatter>
133 import matplotlib.pyplot as plt
134 import pandas as pd
135 from matplotlib import font_manager, rc
136 font_name =
    font_manager.FontProperties(fname="c:/Windows/Fonts/malgun.ttf").get_name()
137 rc('font', family=font_name)
138 df = pd.read_csv('korea.csv',encoding='ms949')
139 plt.figure()
140 plt.scatter(x=df.index,y=df['점수'], marker='2')
141 plt.xticks(range(0,len(df['점수']),1),df['이름'], rotation='vertical')
142 plt.title('학생별 국어점수 산포도')
143 plt.show()
```