

1. 파이썬으로 파일 만들기

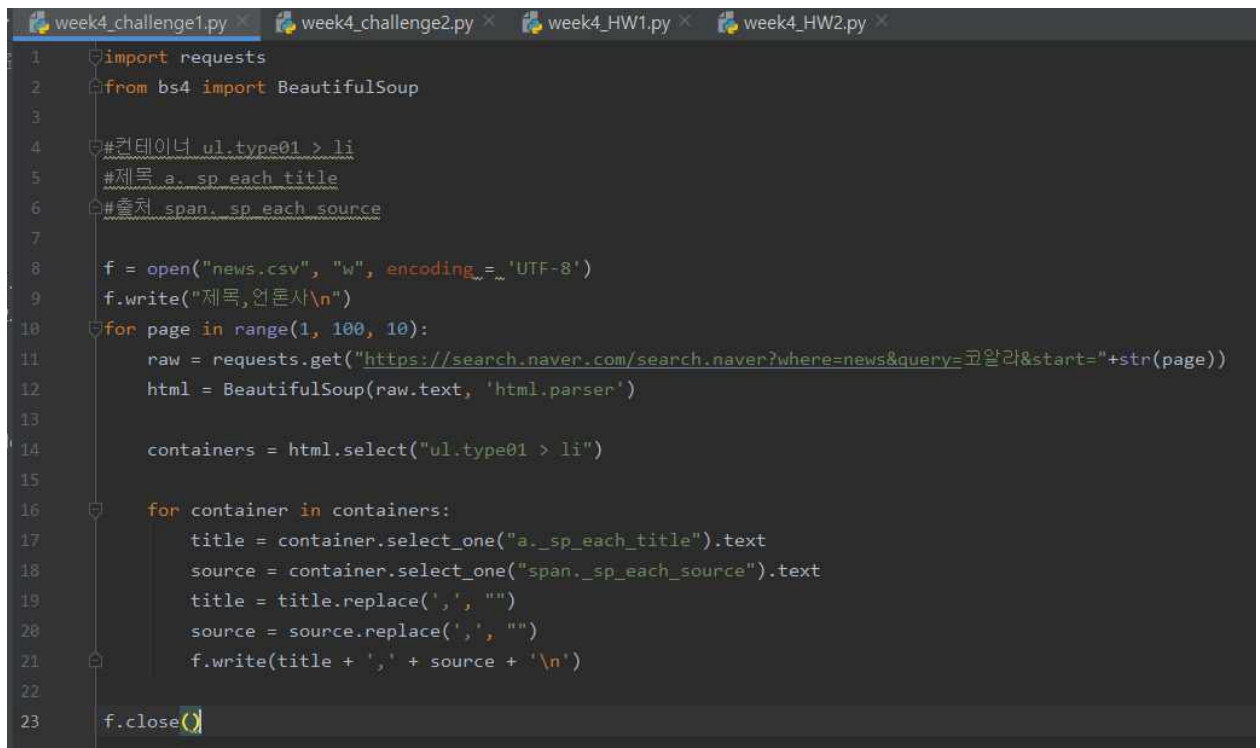
- 콘솔 화면 출력 데이터는 일회성이다.(재사용 위해서는 파일로 변환 필요)
- open함수 -> 실행한 디렉터리에 파일이 생성된다.
 - close() 사용하여 파일 닫기
 - write() 사용하여 파일에 내용 입력
 - ‘a’(append)모드 : 이미 있는 파일 끝부분에 데이터 저장
- 파이썬 with문을 사용하면 close()를 하지 않아도 열고 닫는것을 자동으로 해준다.

```
with open("foo.txt", "w") as f:  
    f.write("Life is too short, you need python")
```

2. 데이터를 csv파일로 저장하기

- CSV : Comma Sparated Value
- 표 위 데이터 분류명(헤더) 사용
- encoding = 'UTF-8'
 - UTF-8이란 유니코드를 인코딩 하는 방식이다
 - 유니코드란 글자와 코드가 컴퓨터가 인식하기 쉽게 1:1 매핑되어 있는 코드표이다.

<challenge 1>



```
1 import requests  
2 from bs4 import BeautifulSoup  
3  
4 #컨테이너 ul.type01 > li  
5 #제목 a._sp_each_title  
6 #출처 span._sp_each_source  
7  
8 f = open("news.csv", "w", encoding='UTF-8')  
9 f.write("제목,언론사\n")  
10 for page in range(1, 100, 10):  
11     raw = requests.get("https://search.naver.com/search.naver?where=news&query=코알라&start="+str(page))  
12     html = BeautifulSoup(raw.text, 'html.parser')  
13  
14     containers = html.select("ul.type01 > li")  
15  
16     for container in containers:  
17         title = container.select_one("a._sp_each_title").text  
18         source = container.select_one("span._sp_each_source").text  
19         title = title.replace(' ', '')  
20         source = source.replace(' ', '')  
21         f.write(title + ',' + source + '\n')  
22  
23 f.close()
```

3. 데이터를 엑셀파일로 저장하기

- Workbook : 엑셀파일 전체 작업공간
- Worksheet : 엑셀 파일의 단일 작업공간
- Cell : 엑셀 파일의 데이터 단위

-wb.active -> 현재 활성화된 sheet 선택

-생성 -> wb.create_sheet()

-시트 제목 변경 -> ws.title = 'title'

-여러개의 시트 중에서 시트 선택 -> ws = wb.get_sheet_by_name('sheet1')
ws = wb['sheet1']

(엑셀은 이전 저장시 마지막에 선택된 시트를 active sheet로 하므로 위의 함수를 사용하는 것이 더 안전)

4. OpenPyXL 심화

-openpyxl.load_workbook("파일이름")

-try, except(예외처리)

<challenge 2>

```
import requests
from bs4 import BeautifulSoup
import openpyxl

# 컨테이너 ul.type01 > li
# 제목 a._sp_each_title
# 출처 span._sp_each_source
keyword = input()
wb = openpyxl.Workbook()
sheet = wb.active
sheet.append(['검색어', '기사제목', '언론사'])

for page in range(1, 100, 10):
    raw = requests.get("https://search.naver.com/search.naver?where=news&query=" + keyword + "&start=" + str(page))
    html = BeautifulSoup(raw.text, 'html.parser')

    containers = html.select("ul.type01 > li")

    for container in containers:
        title = container.select_one("a._sp_each_title").text
        source = container.select_one("span._sp_each_source").text
        title = title.replace(',', '')
        source = source.replace(',', '')
        sheet.append([keyword, title, source])

wb.save("news.xlsx")
```

5. 엑셀 파일 폭

```
for column in sheet.columns:
    max = 0
    column_name = column[0].column_letter
    for cell in column:
        if len(str(cell.value)) > max:
            max = len(str(cell.value))
    width = (max + 2) * 2
    sheet.column_dimensions[column_name].width = width
```