1. **url 변경**

아래의 코드에서 칼부림의 특정 기간을 크롤링하기 위해서는 url을 변경해야합니다.

url에 ds=2023.08.24&de=2023.08.29가 특정기간동안 올라온 기사들을 필터링하는 것입니다. 위 url과 아래 url을 모두 수정해주시면 됩니다.

def makeUrl(search, start\_pg, end\_pg):

    if start\_pg == end\_pg:

        start\_page = makePgNum(start\_pg)

         #https://search.naver.com/search.naver?where=news&query=칼부림&sm=tab\_opt&sort=0&photo=0&field=0&pd=3&ds=2023.07.20&de=2023.08.30&docid=&related=0&mynews=0&office\_type=0&office\_section\_code=0&news\_office\_checked=&nso=so:r,p:from20230720to20230830&is\_sug\_officeid=0&office\_category=0&service\_area=0

        # 정렬을 포함한 url로 수정함.

        url = "https://search.naver.com/search.naver?where=news&sm=tab\_opt&sort=0&photo=0&field=0&pd=3&ds=2023.08.24&de=2023.08.29&docid=&related=0&mynews=0&office\_type=0&office\_section\_code=0&news\_office\_checked=&nso=so:r,p:from20230720to20230830&is\_sug\_officeid=0&office\_category=0&service\_area=0&query=" + search + "&start=" + str(start\_page)

        print("생성url: ", url)

        return url

    else:

        urls = []

        for i in range(start\_pg, end\_pg + 1):

            page = makePgNum(i)

            url = "https://search.naver.com/search.naver?where=news&sm=tab\_opt&sort=0&photo=0&field=0&pd=3&ds=2023.08.24&de=2023.08.29&docid=&related=0&mynews=0&office\_type=0&office\_section\_code=0&news\_office\_checked=&nso=so:r,p:from20230720to20230830&is\_sug\_officeid=0&office\_category=0&service\_area=0&query=" + search + "&start=" + str(

                page)

            urls.append(url)

        print("생성url: ", urls)

        return urls

    # html에서 원하는 속성 추출하는 함수 만들기 (기사, 추출하려는 속성값)

1. **검색어 입력**

코드를 실행한후, 시작페이지와 끝 페이지를 입력하고 검색어 “칼부림”을 입력합니다. 이때 시작페이지와 끝페이지를 입력하는 기준은 실제로 네이버 뉴스에 들어가 해당하는 기간으로 필터링 한뒤에, 끝페이지가 몇 페이지인지 확인한 후 해당 숫자를 끝 페이지 숫자로 입력하면 됩니다.

1. **.csv파일 정리**

얻은 .csv파일을 다른 코드에 작성할 적절한 이름으로 변경해주세요. 이후 이 .csv파일은 댓글을 크롤링할 url의 모음으로서 사용됩니다.