

01. 프로젝트명

파이썬 requests와 bs4를 이용한 정적인 페이지 크롤링

02. 프로젝트 설명 (200자)

네이버 지식인 페이지의 정보를 원하는 형태로 추출하고 엑셀에 저장한다.

03. 프로젝트 과제 상세 : 프로젝트를 진행할 때, 유념해야할 것들입니다.

기준	과제 상세
메인페이지 크롤링	다음 조건에 해당하는 페이지 URL을 찾고 requests 요청을 보냅니다.
	검색어 : 반도체
	기간 : 2023년 8월 1일 ~ 2023년 8월 31일
	정렬 : 최신순
상세페이지 크롤링	HTML 구조를 분석해 나무태그를 찾아주고 CSS선택자를 만들어 줍니다.
	1페이지 10개 질문에 대해서 날짜, 제목, 링크, 답변수를 가져와서 출력합니다. (출력 형식은 아래 첫번째 스크린샷과 동일하게)
	질문 상세페이지 링크로 requests 요청을 보내고 새로운 soup를 만들어 줍니다.
	상세페이지에서 조회수를 가져와서 출력합니다. (출력 형식은 아래 두번째 스크린샷과 동일하게)
데이터 저장	2페이지까지 크롤링할 수 있도록 코드를 수정합니다.
	크롤링할 데이터를 저장할 2차원 리스트를 만들어 줍니다.
	제목, 링크, 날짜, 답변수, 조회수 컬럼을 가지는 데이터프레임을 만들어 줍니다. (답변수와 조회수는 숫자형태로)
	데이터프레임을 'result.xlsx' 파일로 저장합니다. (파일 형태는 아래 세번째 스크린샷과 동일하게)

04. Project Material : 프로젝트를 진행할 때, 필요한 정보입니다. Starter 를 위한 기본 코드도 이에 해당합니다.

<메인페이지 크롤링 후 출력 결과>

```
2023.08.31. 대한민국 경제 급성장 이유 https://kin.naver.com/qna/detail.naver?d1id=4&dirId=408&docId=453685135&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=1&search_sort=3&spq=0 2
2023.08.31. 구리 인창고vs중화고 https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110308&docId=453582509&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=2&search_sort=3&spq=0 1
2023.08.31. 진로 결정 https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110404&docId=453671356&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=3&search_sort=3&spq=0 1
2023.08.31. 내신 5등급 성균관대 https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110408&docId=453668459&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=4&search_sort=3&spq=0 2
```

<상세페이지 크롤링 후 출력 결과>

```
2023.08.31. 대한민국 경제 급성장 이유 https://kin.naver.com/qna/detail.naver?d1id=4&dirId=408&docId=453685135&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=1&search_sort=3&spq=0 2 72
2023.08.31. 구리 인창고vs중화고 https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110308&docId=453582509&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=2&search_sort=3&spq=0 1 155
2023.08.31. 진로 결정 https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110404&docId=453671356&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=3&search_sort=3&spq=0 1 15
2023.08.31. 내신 5등급 성균관대 https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110408&docId=453668459&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=4&search_sort=3&spq=0 2 60
```

<데이터저장 후 엑셀 파일>

	A	B	C	D	E	F
1		날짜	제목	링크	답변수	조회수
2	0	2023.08.31.	대한민국 경제 급성장 이유	https://kin.naver.com/qna/detail.naver?d1id=4&dirId=408&docId=453685135&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=1&search_sort=3&spq=0	2	72
3	1	2023.08.31.	구리 인창고vs중화고	https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110308&docId=453582509&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=2&search_sort=3&spq=0	1	156
4	2	2023.08.31.	진로 결정	https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110404&docId=453671356&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=3&search_sort=3&spq=0	1	15
5	3	2023.08.31.	내신 5등급 성균관대	https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110408&docId=453668459&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=4&search_sort=3&spq=0	2	60
6	4	2023.08.31.	적정기술의 사례와 사용된 과학용어	https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110412&docId=453671356&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=3&search_sort=3&spq=0	1	434
7	5	2023.08.31.	다이오드 질문입니다	https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110416&docId=453671356&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=3&search_sort=3&spq=0	2	22
8	6	2023.08.31.	중국 반도체 굴기 해결방안	https://kin.naver.com/qna/detail.naver?d1id=4&dirId=408&docId=453685135&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=1&search_sort=3&spq=0	2	124
9	7	2023.08.31.	성균관대 시스템반도체	https://kin.naver.com/qna/detail.naver?d1id=11&dirId=110408&docId=453668459&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=4&search_sort=3&spq=0	4	104
10	8	2023.08.31.	태양광 발전기 효율적이게 설치,	https://kin.naver.com/qna/detail.naver?d1id=5&dirId=508&docId=453671356&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=3&search_sort=3&spq=0	4	64
11	9	2023.08.31.	대기업 반도체 생산직 질문	https://kin.naver.com/qna/detail.naver?d1id=4&dirId=408&docId=453685135&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=1&search_sort=3&spq=0	1	39
12	10	2023.08.31.	2021 k8 월석 스마트 주차보조 중	https://kin.naver.com/qna/detail.naver?d1id=8&dirId=808&docId=453671356&q=67CY64+E7LK0&enc=utf8&ion=kin&rank=3&search_sort=3&spq=0	1	202