

본 자료 및 영상 콘텐츠는 저작권법 제25조 2항에 의해 보호를 받습니다. 본 콘텐츠 및 콘텐츠 일부 문구등을 외부에 공개, 게시하는 것을 금지합니다. 특히 자료에 대해서는 저작권법을 엄격하게 적용하겠습니다.

Scrapy : 크롤링 프레임워크 연습

사전 작업

1. Scrapy 프로젝트 구성 및 주요 파일 확인

```
> scrapy startproject mycrawler(프로젝트이름)
> cd mycrawler
> scrapy genspider test_web(spider이름) davelee-fun.github.io(사이트주소)
```

2. Scrapy 테스트

```
> scrapy shell davelee-fun.github.io
> view(response)
> response.url
```

3. 쉘에서 실행해보기 (문제1)

- All Stories 의 추천 타이틀 리스트로 출력해보기
- All Stories 의 상품명 리스트로 출력해보기

문제1 퀴즈 정답

3. 쉘에서 실행해보기 (문제)

- All Stories 의 추천 타이틀 리스트로 출력해보기
- All Stories 의 상품명 리스트로 출력해보기

```
> response.css('section.recent-posts div.card-body > h2 > a::text').getall()  
> response.css('section.recent-posts div.card-body > h4::text').getall()
```

4. scrapy 프로젝트로 추천 타이틀과 상품명 리스트로 json 파일 만들어보기 (문제2)

문제2 퀴즈 정답

4. scrapy 프로젝트로 추천 타이틀과 상품명 리스트로 json 파일 만들어보기 (문제2)

- [items.py](#) 업데이트
- [test_web.py](#) 업데이트
- [settings.py](#) 업데이트
- 다음 쉘명령으로 crawling 및 json 파일 저장

```
scrapy crawl test_web -o test_web.json -t json
```

각 업데이트 코드 파일은 별도 프로젝트 파일로 제공함