웹 크롤링 구조

https://~~ =>이 주소는 사용자가 해당주소로 정보를 get post delete 등의 행위를 하기위해, IP주소대신에 변환된 형태로 사용하는것

요청(get post등)을 받았을경우 HTML파일을 송신해준다.

이렇게 가져온 HTML파일을 파싱한다.

(HTML파일,parser\_name)

으로 객체를 생성한뒤

get\_find로 HTML파싱부분중 필요부분을 get\_text로 가져오는 느낌

HTML구조

태그 이외에 속성을 가지고 필요부분을 추출하는 방법이 중요 ->

Css 프로퍼티 별 내용자르기

