

오늘은 여러 페이지에 있는 테이블 데이터를 스크래핑하여 csv 파일로 만들어주는 프로그램을 만들어 보았다.
웹 브라우저에서 흔히 볼 수 있는 표 형태의 리스트, 커뮤니티 게시판 데이터를 엑셀 파일로 만드는 프로그램이라고 보면 된다.

프로그램을 사용하고자 하는 사이트가 페이지 접근을 위해서는 로그인이 필요하고, 한 페이지가 아닌 여러 페이지에 테이블 데이터가 필요한 상황이었다.
향후 페이지가 추가되더라도 정상 작동하도록 구현하였다.

코드 작성하는 데 대략 3시간 정도 썼다.
수준 높은 코드는 아니라고 생각한다.
대신 그만큼 어렵지 않고 짧다(76 lines).
파이썬 selenium 프레임워크를 처음 이용하거나 스크래핑한 데이터를 csv 파일을 만드는 방법에 익숙지 않은 사람이 참고하면 좋을 듯.

requests, bs4를 이용하려고 했는데 영 진전이 없어서 그냥 쉬운 방식을 선택했다.
두 녀석은 다음에 다시 도전해 봐야겠다.
코드를 업로드한 깃허브 페이지 링크를 아래 적어놓았다.
<https://github.com/c2lv/ListScrapingProgram>

