1. 코스피 / 코스닥 종목 가져와 출력하기

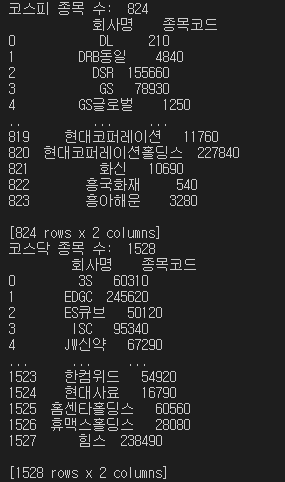
* code sample

def get\_item\_kospi(self):

        df\_stockMkt = pd.read\_excel(os.path.join(BASE\_DIR, '상장법인목록\_stockMkt.xlsx'))

        print('코스피 종목 수: ', len(df\_stockMkt))

        print(df\_stockMkt[['회사명', '종목코드']])

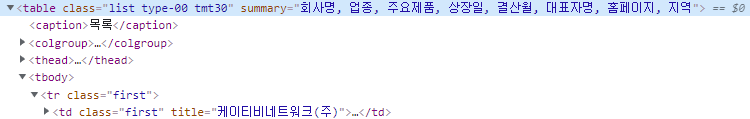
  
[이슈]  
1. pd.read\_excel로 xls파일을 읽을 수 없는 문제  
 해결: xls파일을 xlsx로 변환 (<https://blog.naver.com/jin31200/221355318078>)  
2. Pandas 기본 엔진이었던 xlrd가 업데이트되면서 xlsx를 지원하지 않는 문제  
 해결: Pandas 기본 엔진을 openpyxl로 변경 (pip install openpyxl)  
 (<https://steadiness-dev-invest.tistory.com/141>)

1. 심화과정(크롤링)

**사용 기술**

* Python
* Selenium
* BeaurifulSoup
* ChromeDriver
* csv

**구현 방법**

* 1. 웹 개발자 도구에서 class name을 활용해 테이블을 구성하는 element에 접근  
     - 웹 개발자 도구  
       
     - code sample

soup = BeautifulSoup(browser.page\_source, "html.parser")

data\_rows = soup.find("table", "list type-00tmt30")

.find("tbody").find\_all("tr")

* 1. 각 row마다 element의 text값을 가져와 리스트 형태로 전달

for row in data\_rows:

        columns = row.find\_all("td")

        # 현재 행의 text값을 가져온다

        data = [column.get\_text().strip() for column in columns]

* 1. 리스트를 넘겨받아 csv파일에 기록

now = datetime.datetime.now().strftime('%Y-%m-%d %H %M')

filename = BASE\_DIR + '/상장법인목록(' + now + ').csv'

f = open(filename, 'w', encoding='utf-8-sig', newline='')

writer = csv.writer(f)

...

writer.writerow(data)

* 1. 다음 페이지 버튼으로 페이지를 이동하고 1~3을 반복 (마지막 페이지까지 도달하면 종료)

while True:

...

curr\_page = soup.find("div", "info type-00").find("strong")

.get\_text().split()

    # 마지막 페이지면 종료

    if int(curr\_page[0]) == int(full\_page):

        break;

# 다음 페이지

    element = browser.find\_element\_by\_class\_name("next")

    browser.execute\_script("arguments[0].click();", element)