데이터 크롤링과 정제

과제3: 네이버 증시 정보 크롤링

과제 3-1

- ■시가 총액 10위까지의 기업 정보를 크롤링
 - 네이버 금융 웹사이트
 - https://finance.naver.com/sise/sise_market_sum.naver
 - 크롤링 항목 7개 출력
 - 종목명, 종목코드, 현재가, 전일가, 시가, 고가, 저가



- 네이버 웹사이트의 한글 깨짐 문제
 - requests 라이브러리 사용

```
import requests

html = requests.get(url)
soup = BeautifulSoup(html.text, 'html.parser')
```

네이버 주식: 시가 총액현황 #1

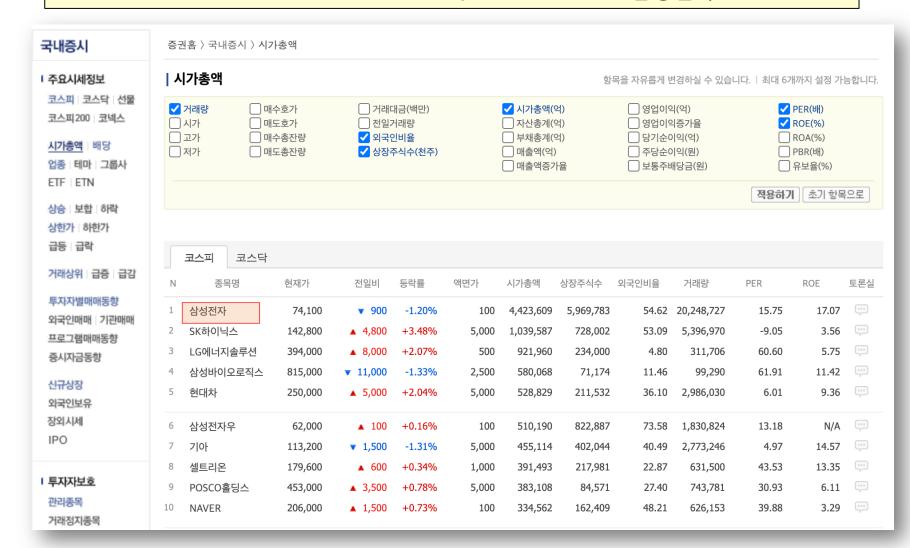
- 왼쪽 "시가 총액"메뉴에서 상위 10개 기업 정보 크롤링
 - 시가 총액 url: https://finance.naver.com/sise/sise_market_sum.naver

국내증시 증권홈 〉국내증시 〉시가총액 │ 시가총액 주요시세정보 항목을 자유롭게 변경하실 수 있습니다. | 최대 6개까지 설정 가능합니다. 코스피 코스닥 선물 ✓ 거래량 매수호가 거래대금(백만) ✓ 시가총액(억) 영업이익(억) V PER(배) 코스피200 코넥스 시가 메도호가 전일거래량 자산총계(억) 영업이익증가율 ✓ ROE(%) 고가 매수총잔량 ✓ 외국인비율 부채총계(억) 당기순이익(억) ROA(%) 시가총액 배당 저가 ✓ 상장주식수(천주) 매도총잔량 매출액(억) 주당순이익(원) PBR(배) 업종 테마 그룹사 매출액증가율 보통주배당금(원) 유보율(%) ETF ETN **적용하기** 초기 항목으로 상승 보합 하락 상한가 하한가 급등 급락 코스피 코스닥 거래상위 급증 급감 종목명 현재가 전일비 등락률 액면가 시가총액 상장주식수 외국인비율 거래량 PER ROE 토론실 투자자별매매동향 삼성전자 74,100 **900** -1.20% 4,423,609 5,969,783 54.62 20,248,727 15.75 17.07 외국인매매 기관매매 5,396,970 2 SK하이닉스 **4,800** +3.48% 1,039,587 728,002 -9.05 3.56 142,800 5,000 53.09 프로그램매매동향 LG에너지솔루션 394,000 **&** 8,000 +2.07% 921,960 234,000 4.80 60.60 5.75 311,706 증시자금동향 삼성바이오로직스 815,000 **v** 11,000 -1.33% 2,500 580,068 71,174 11.46 99,290 61.91 11.42 신규상장 현대차 211,532 250,000 **5,000** +2.04% 5,000 528,829 36.10 2,986,030 6.01 9.36 외국인보유 장외시세 삼성전자우 62,000 **100** +0.16% 100 510,190 822,887 73.58 1,830,824 13.18 N/A IPO **v** 1,500 402,044 14.57 기아 113,200 -1.31% 5,000 455,114 40.49 2,773,246 4.97 셀트리온 179,600 **4** 600 +0.34% 1,000 391,493 217,981 22.87 631,500 43.53 13.35 투자자보호 POSCO홀딩스 453,000 **3,500** +0.78% 383,108 84,571 27.40 743,781 30.93 6.11 5,000 관리종목 10 NAVER 206,000 **1,500** +0.73% 100 334,562 162,409 48.21 626,153 39.88 3.29 거래정지종목

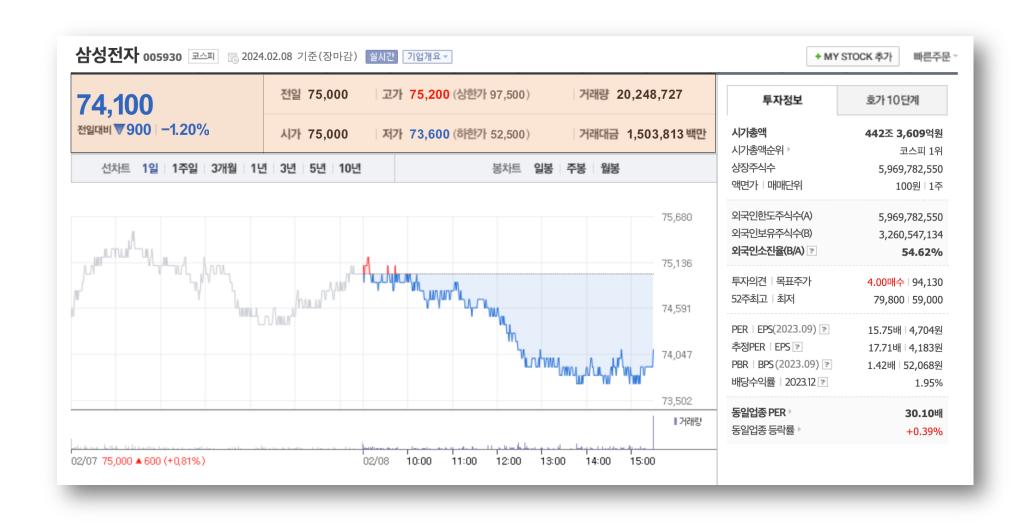
네이버 주식: 시가 총액현황 #2

■ 각 회사별 URL 및 이름을 리스트에 저장

삼성전자



삼성전자 주식 세부 현황 페이지



프로그램 동작 및 실행 결과

- ■프로그램 동작
 - 코스피 상위 10개 기업의 URL 및 기업 이름을 리스트에 저장
 - 메뉴에서 선택한 기업의 세부 주식 정보를 화면에 출력
 - 사용자가 -1을 입력할 때 까지 계속 반복함

```
[ 네이버 코스피 상위 10대 기업 목록 ]
[ 1] 삼성전자
[2] SK하이닉스
[ 3] LG에너지솔루션
[ 4] 삼성바이오로직스
[5] 현대차
[6] 삼성전자우
[7] 기아
[8] 셀트리온
[ 9] POSCO홀딩스
[10] NAVER
주가를 검색할 기업의 번호를 입력하세요(-1: 종료): 1
https://finance.naver.com/item/main.naver?code=005930
종목명: 삼성전자
종목코드: 005930
현재가: 74,100
전일가: 75,000
시가: 75,000
고가: 75,200
저가: 73,600
```

실행 결과



Questions?