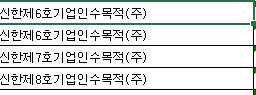
종목선정 데이터 처리 순서

1. 202006 이후 상장 삭제 ⇒ 2019년 이전 상장 삭제 ,상장일 칼럼 삭제(O)
2. spac 제거 .



⇒ ‘기업인수목적’ 키워드로 삭제(O)

1. 재무제표부재 기업 삭제 (O)

⇒ ‘매출액증가율(IFRS)’= na and ’순이익증가율(IFRS)’ =na 면 삭제

1. 자본(\*)(IFRS) 열 삭제(O)
2. 자본금 열 추가 (우선주자본금(IFRS)+보통주자본금(IFRS))(O)
3. 차입금의존도 na값 0으로 채우기 (O)
4. 영업이익 + 순이익 으로 순이익 증가율 0 처리

⇒ 기준 마련 필요.

1. 순이익 증가율이 0이고 당기에도 순손실 기록시 삭제.
2. 영업이익이 전기대비 증가 + 흑자전환시 보류
3. 행이름 수정(완)(O)

회사명 = company\_name

거래소코드 = ticker

회계년도 = fiscal\_year

매출액증가율(IFRS) = sales\_growth\_rate

순이익증가율(IFRS) = net\_profit\_growth\_rate

매출액정상영업이익률(IFRS) = operating\_profit\_rate

자기자본순이익률(IFRS) = roe

총자본순이익률(IFRS) = roa

부채비율(IFRS) = debt\_ratio

차입금의존도(IFRS) = tbp(Total Borrowings And Bonds Payable to Total Assets)

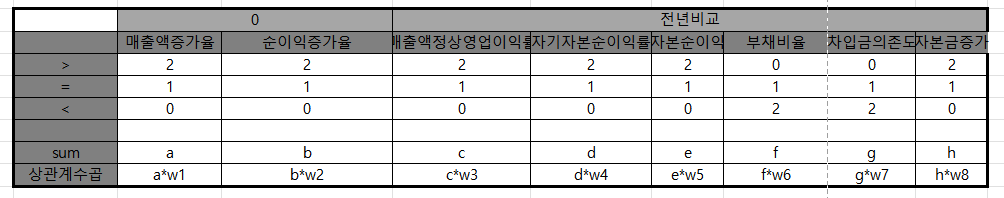
자본금(\*)(IFRS) = equity\_capital

\* (정상)영업손익(보고서기재)(IFRS) = operating\_profit

당기순이익(손실)(IFRS) = net\_profit

상장일 = listing\_date

주가 수익률 = Close\_ratio

1. 시가총액 200억 미만 삭제 (O 위에 하면서 다 사라짐)
2. 2021년 반기보고서 기준 순손실(순이익란 -)인 기업 삭제(적용안함)
3. 상장폐지 제거 && 현재 거래정지 제거 (O)
4. 상관분석을 위한 Y(O)
5. 작년 순손실기록 기업의 경우 흑자전환하였다고 하더라도 순이익증가율 0으로 처리()
6. 펀더멘탈 저조한 기업 처리 ⇒ 최근 12개월 주가수익률이 높은 종목 상위 10% 생존 (패자부활전)
7. 업종분류: 주식시장 업종 분류 (WICS), 한국표준산업분류(포기)

⇒ wics는 다트나 krx에서 안뽑힘….

⇒ 한국표준산업분류는 대분류도 20개가 넘어감

csv 변경사항

1. 회계년도 2021 되어 있는것 17개 삭제
2. 회계년도 yyyy형식으로 변환
3. 외국 회사 자본금 없는 회사 삭제
4. 에스브이인베스트먼트 삭제
5. 밑에 붙인것들 , 제거

python 진행 사항(project1)

1. 회계년도 19년 이전 삭제
2. 상장일 20200601 작은거 삭제 상장일 제거
3. 스팩주 삭제
4. tbp 차입금의존도 na -> 0
5. sales\_growth\_rate 빈 값들 삭제
6. 20, 21 년 데이터 하나라도 없는 회사는 삭제
7. 시가총액 자르기(200억 미만)
8. 회계년도 2020,2021 합치기 (34개 회사 **Close\_ratio** NAN값
9. 점수로 변환
10. 히트맵 그리기(상관계수)

python 진행 사항(project1\_daily)

1. pykrx 이용해서 일별 데이터 받아오기

(**시가,고가,저가,거래량,거래대금,등락률,PBR,기관합계,기타법인,개인,외국인합계,공매도,공매도매수,비중,거래대금/시가총액)**

**'시가','고가','저가','거래량','거래대금','등락률','PBR','기관합계','기타법인','개인','외국인합계','공매도',"공매도매수",'거래대금/시가총액','종가'**

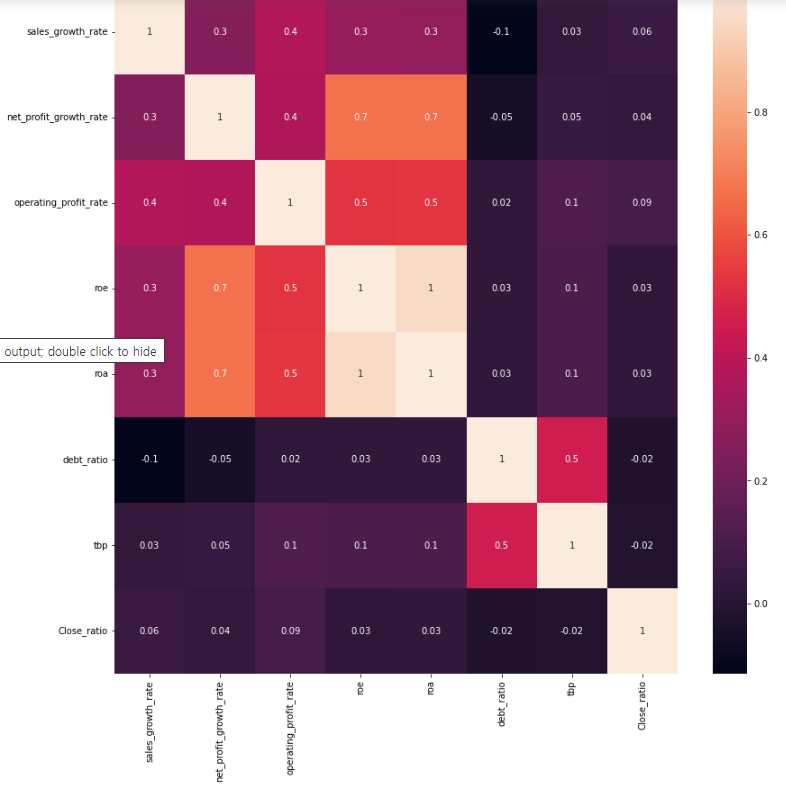
증권거래소에서는 상장종목을 기업규모에 따라 대·중·소형주로 분류하여 주가지수를 산출함으로써 각종 시황분석 및 포트폴리오 성과측정지표로 활용하도록 하고 있다.

1983년부터 자본금규모별 주가지수를 산정하였으나 실제 시장상황과 주가지수와의 괴리가 발생하고 주요 금융선진국에서는 대부분 기업규모별 지수를 시가총액으로 분류하고 있어 2003년부터 2000년 1월 4일을 1,000포인트로 하는, 시가총액규모별 주가지수를 산출, 발표하고 있다.

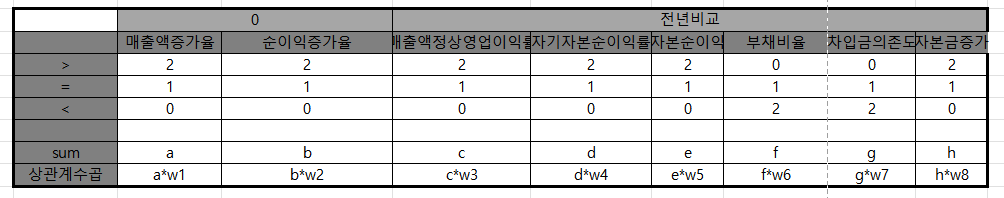
산출대상종목은 거래소 전 상장종목을 대상을 하되 증권투자회사, 관리종목, 정리매매종목 등은 제되된다.

기업규모 분류기준은 정기변경일 이전 3개월간의 일평균 시가총액을 기준으로 하며, 시가총액 상위 100위까지를 '대형주'로 하고 상위 101위부터 300위까지를 '중형주'로, 나머지 종목을 '소형주'로 분류한다.

또 지수의 대표성 제고 및 정확한 시장상황 반영을 위해 매년 2차례 지수선물 및 지수옵션시장의 6월물과 12월물의 최종 거래일의 다음 매매거래일에 구성종목을 교체한다.

****

매출액 증가율 / 순이익 증가율 / 영업이익률 / 자기자본이익률 / 총자산이익률 / 부채비율 / 차입금의존도

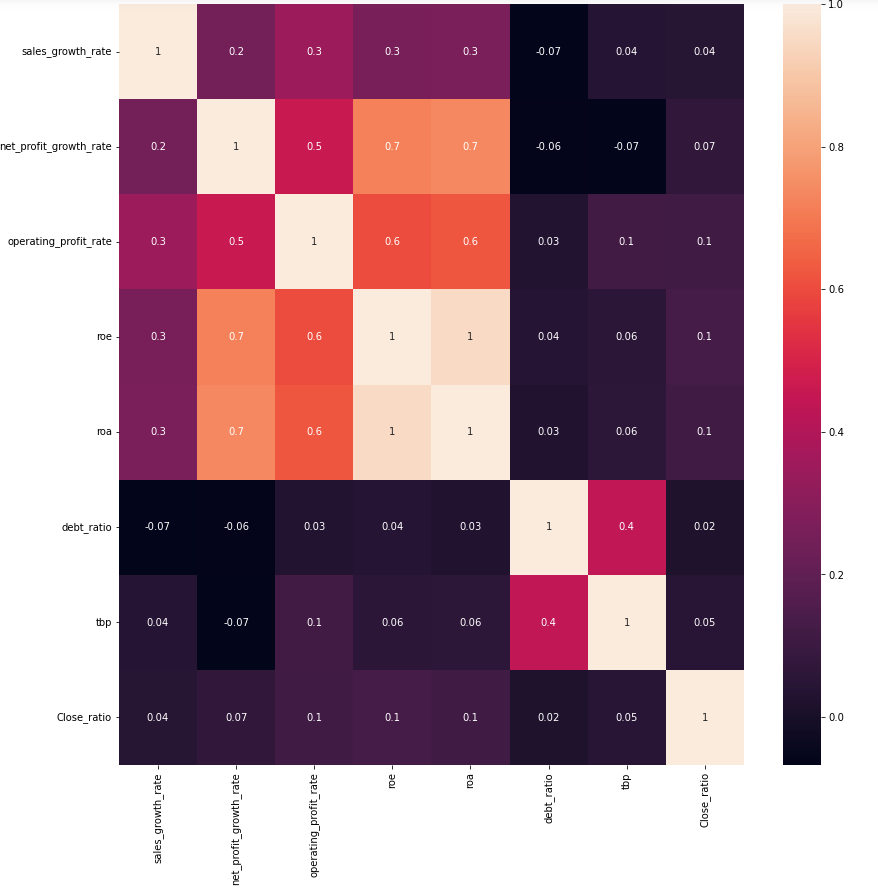


~~우리나라 시장은 원시상태 (비합리적) -실적을 보지 않음.?~~

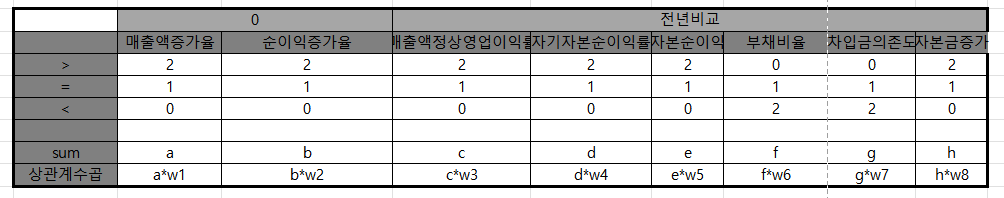
⇒ 전체 종목을 갖고 상관분석을 할 경우 과도하게 다양한 기업들이 섞여 있어 특정 피처의 영향을 확인이 불가능하게 된다고 해석

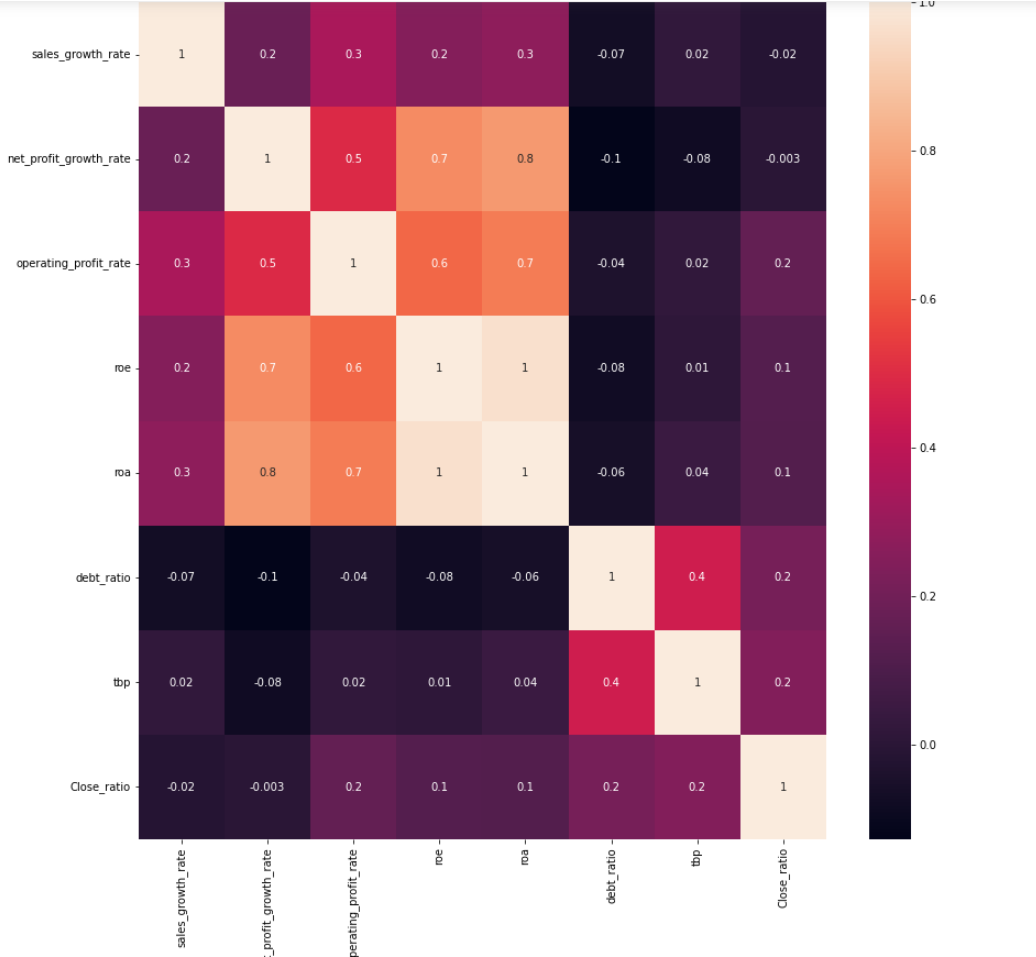
자본금의 영향이 가장 큰 것은 코로나 동안 대형주위주의 흐름이 강했던 영향?

유의확율을 볼 것도 없이 상관계수가 너무 낮음…



5000억 미만



3조 이상

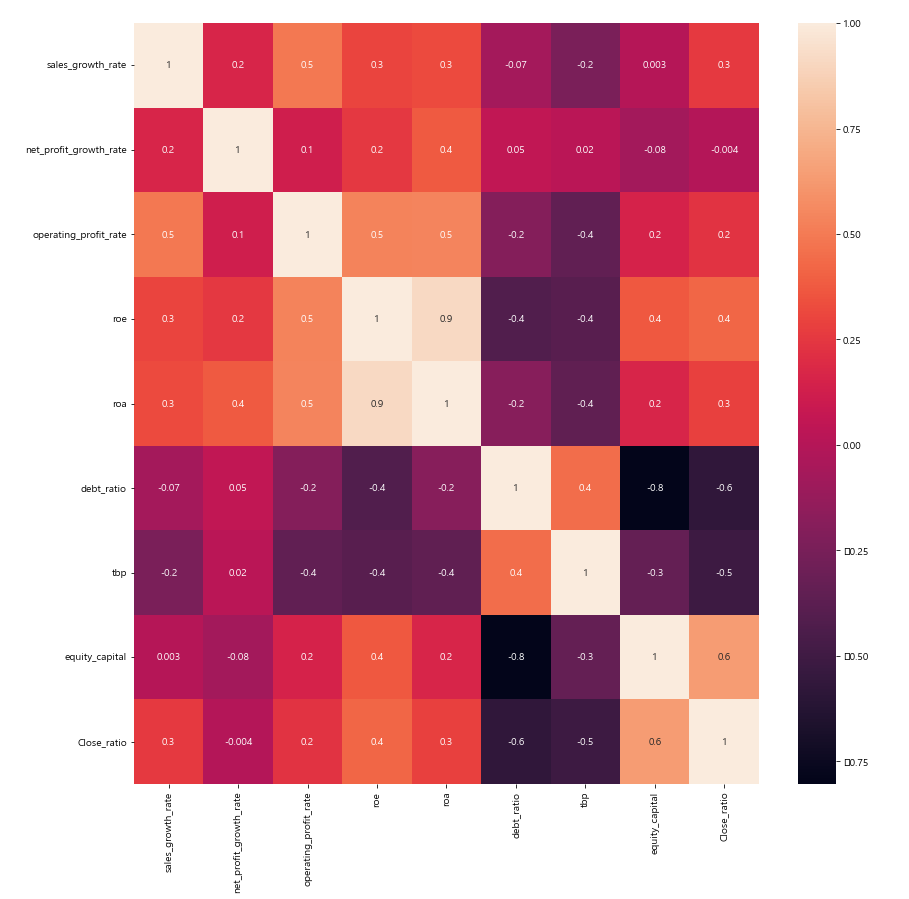
⇒ 3조이상과 5000천억 미만 상관관계를 통해 수익성이 주가상승률에 큰 영향을 미친다는 것을 알 수 있었고, 5천억미만기업의 경우 안정성(부채비율, 차입금의존도)의 영향이 크지 않으나 3조 이상 기업의 경우 안정성에 따라 주가상승률에 영향이 생긴다는 것을 알 수 있었다.

상관관계 변수들 사이의 관계고민이 필요함

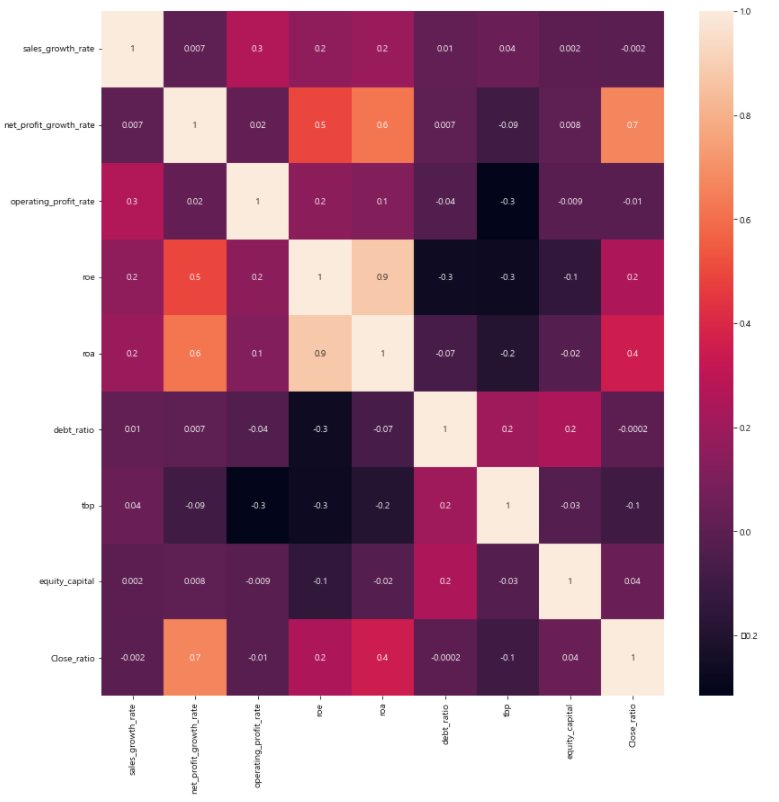
전체는 힘들것으로 예상되나, 정규분포로 만들어서 돌려보기로함.

**<시가총액기준 대형주/중형주/소형주 별 재무재표vs주가상승률 상관분석 결과>**

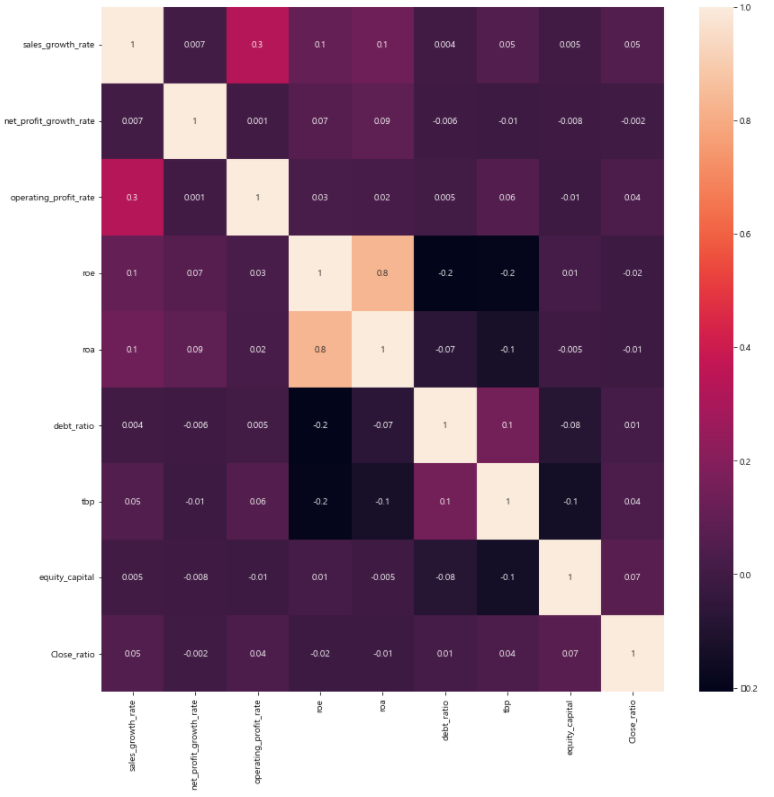
1. **대형주 - 시가총액 기준 : 상위 100위 이상**
2. **중형주 - 시가총액 기준 : 상위 101 ~ 300위**
3. **소형주 - 시가총액 기준 : 상위 301위 이하**
4. **대형주**

****

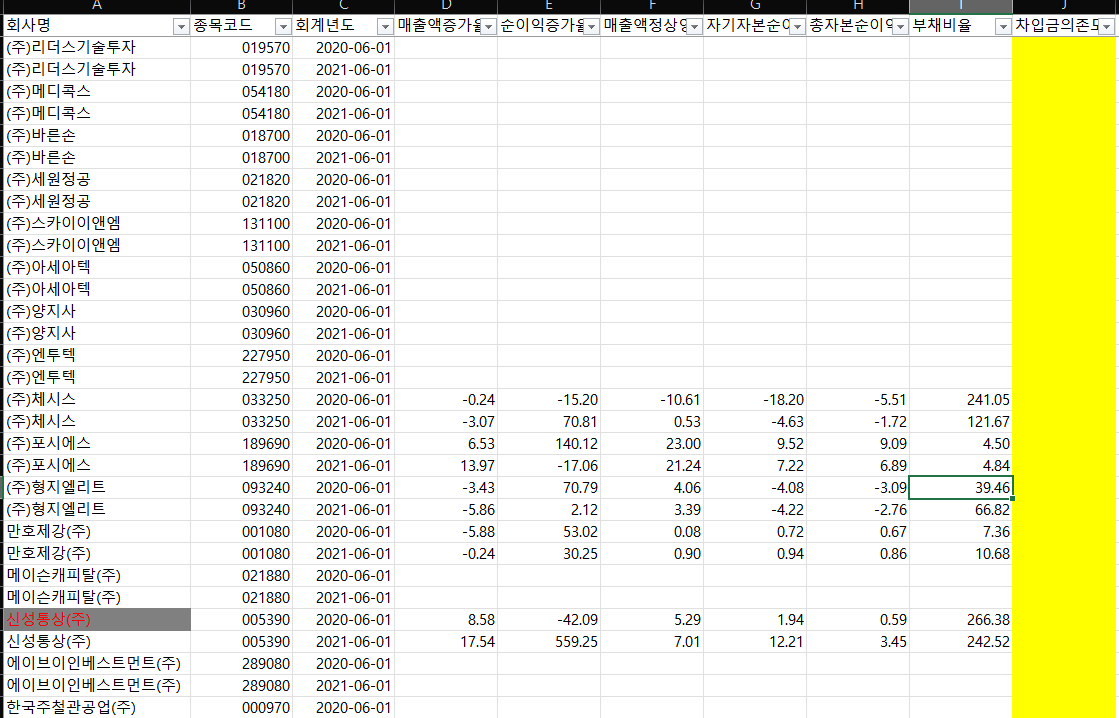
1. **중형주**

****

1. **소형주**

****

**이슈**

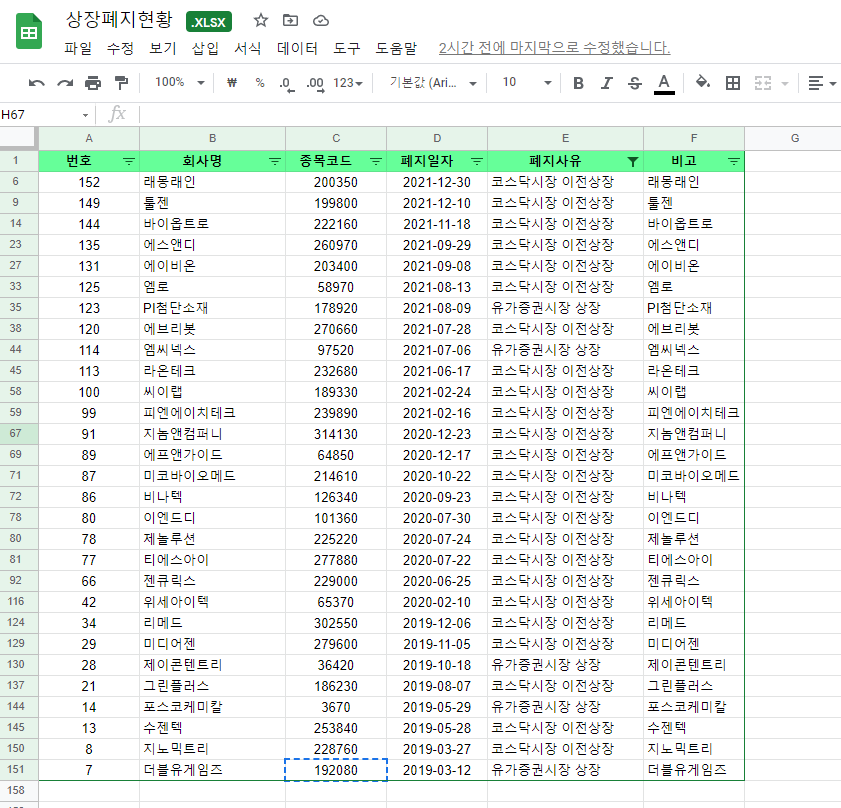


빈칸 직접 dart에서 조회하여 채움

거래정기 기간이 장기일 경우

* 거래정지 기간중 가격은 거래직전가격으로 입력

수정주가 2021년0630일로 계산



상장폐지 예외 종목