

코로나 19

소비데이터 분석을 통한 주가 예측

2020 금융 빅데이터 챌린지



말티즈는 참지않기

Contents

1

소비데이터 분석

- 업종별 매출액 비교
- 업종별 거래량 비교

2

키워드 분석

- 채널 별 키워드 분석
- 연관 키워드 분석

3

주식 종목 분석

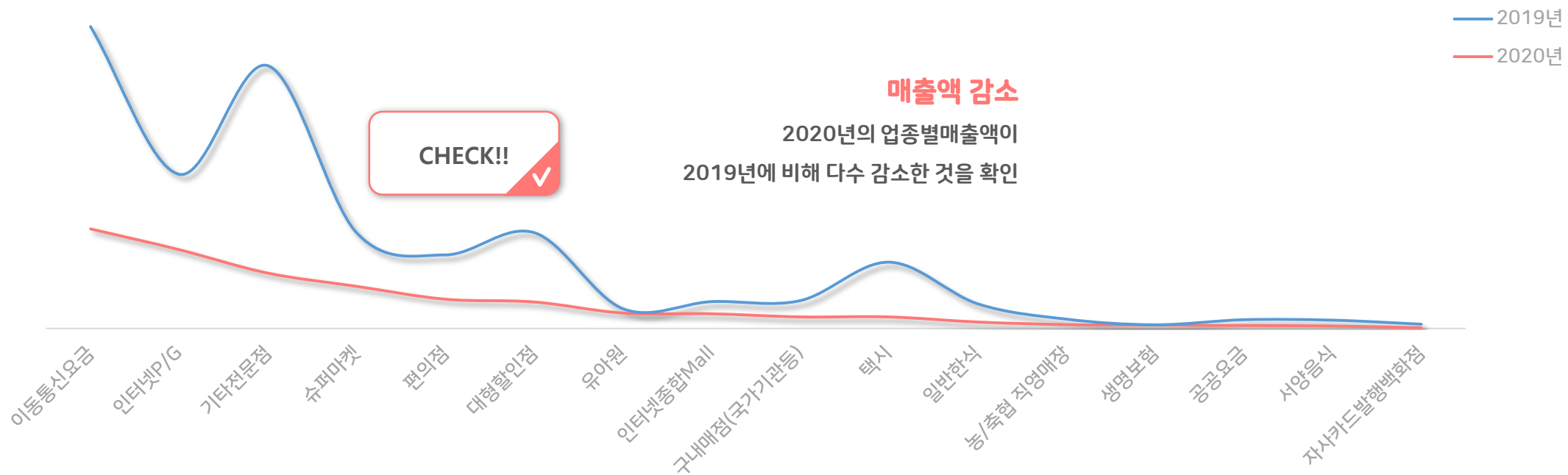
- 소비데이터를 활용한 분석
- 주식키워드 분석
- 키워드 관련 종목

4

결론

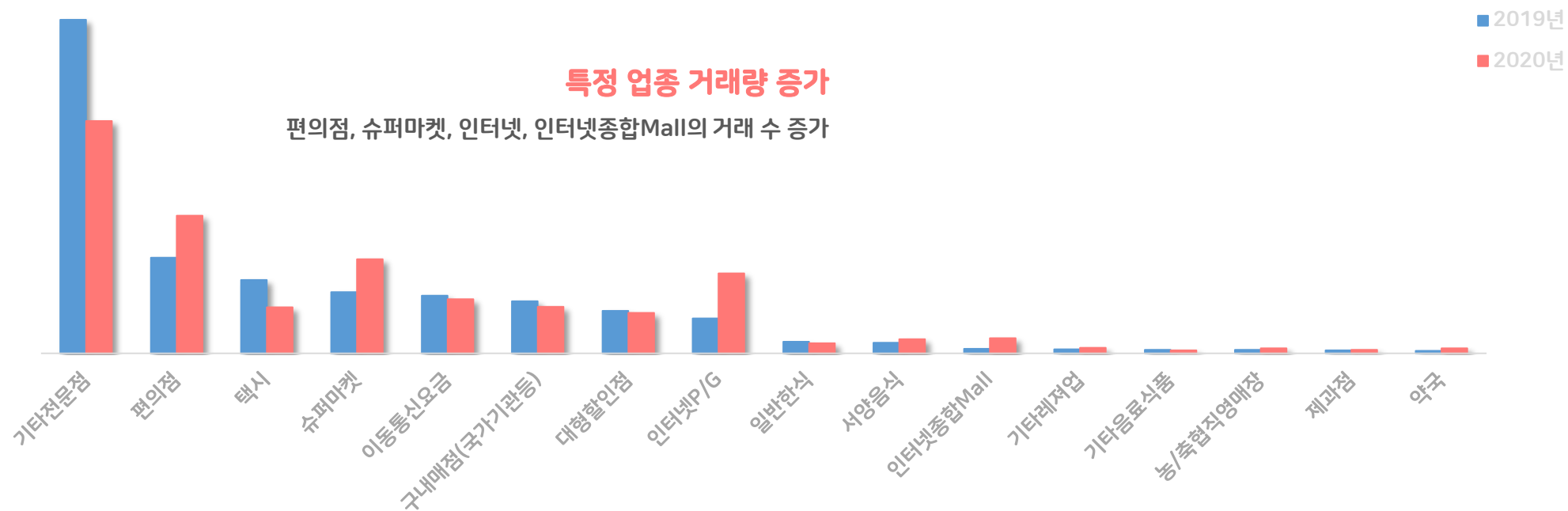
업종별 매출액 비교

업종 품목별 소비데이터 분석



업종별 거래량 비교

업종 품목별 소비데이터 분석



채널 별 키워드 확인

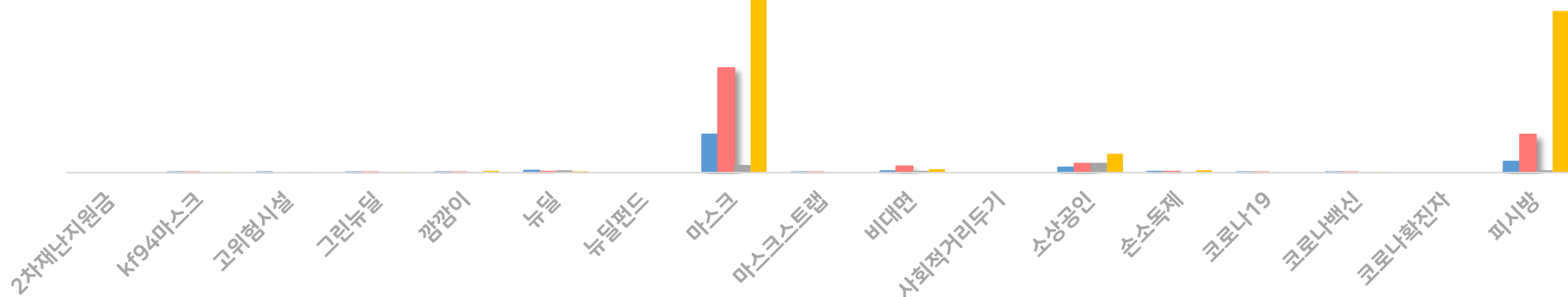
2019년도 키워드 버즈 추이

총 개수 1642804

ko.blog
ko.insta
ko.news
ko.twitter

2019년 채널 별 키워드 값

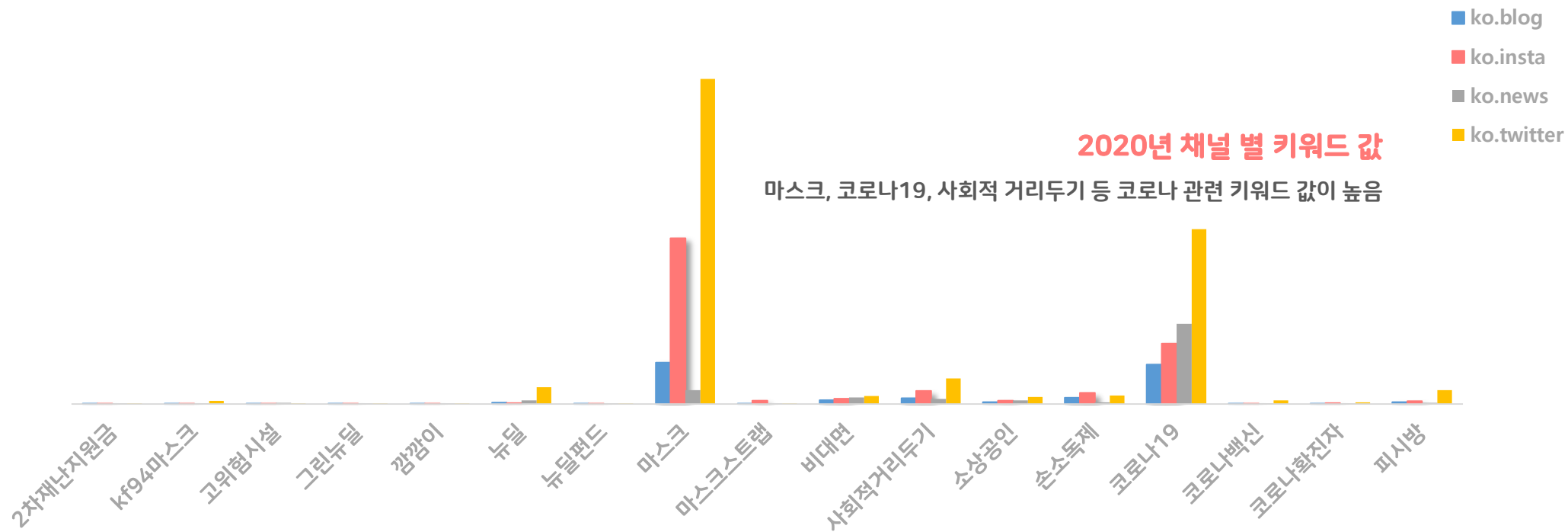
비대면, 마스크, 소상공인, 피시방 키워드 값이 높음



채널 별 키워드 확인

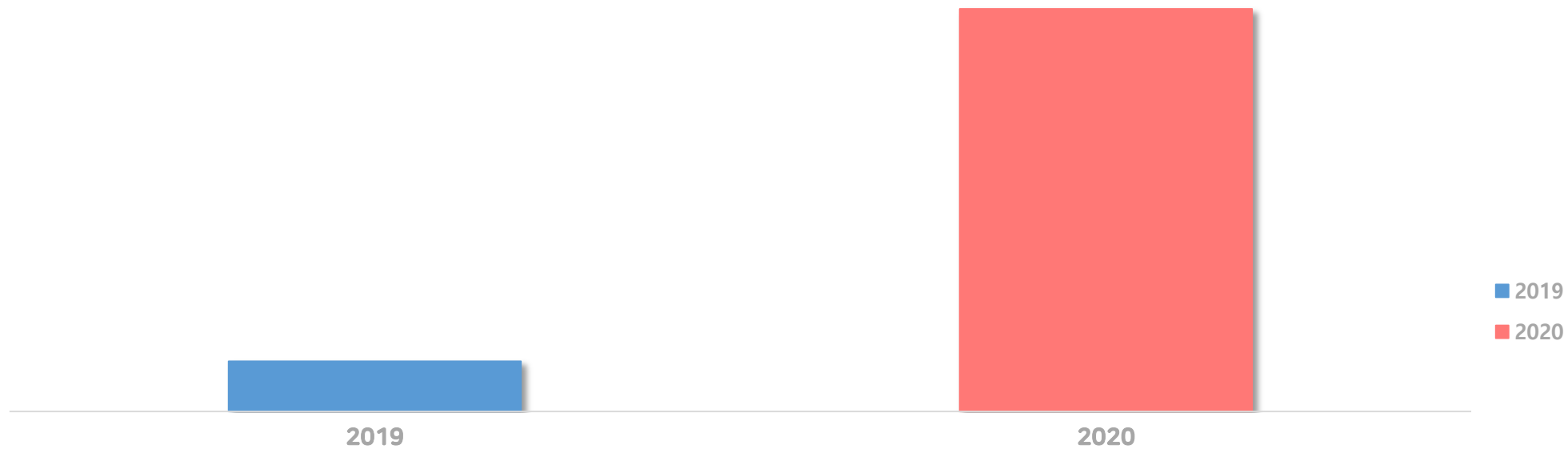
2019년도 키워드 버즈 추이

총 개수 13238450



채널 별 키워드 확인

2019년과 2020년의 비교

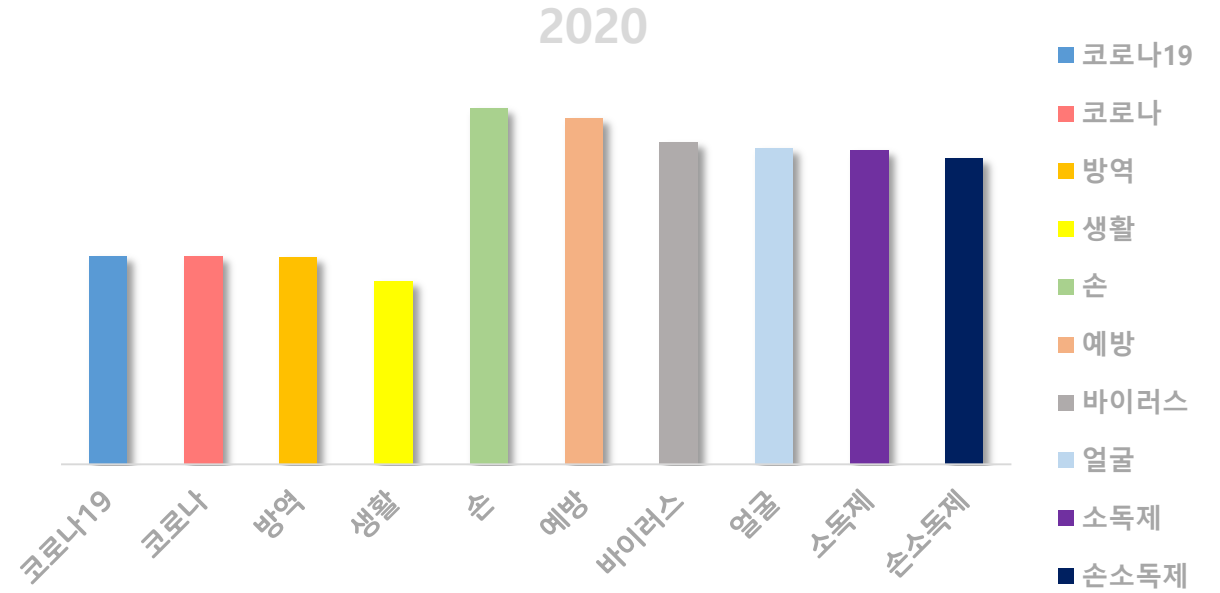
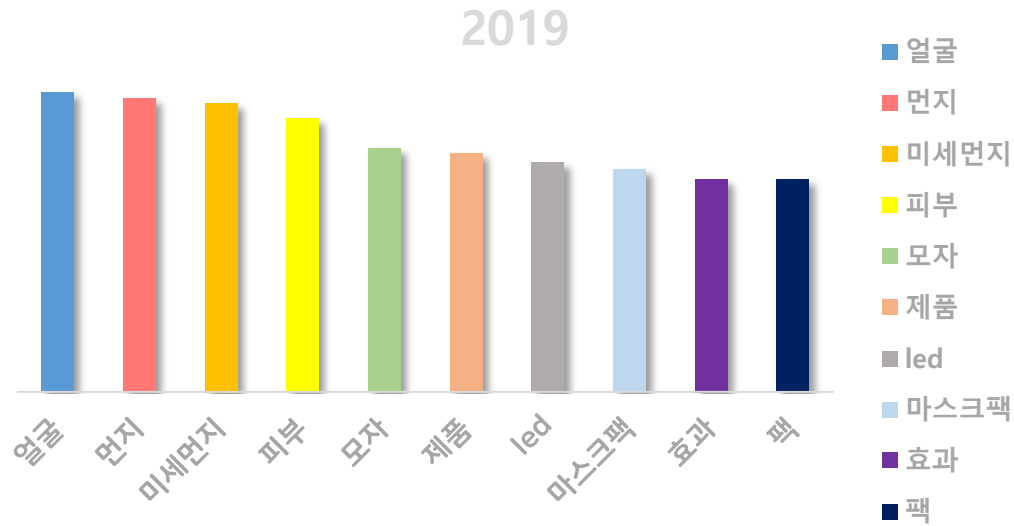


2019년 2020년 키워드 수 비교

2020년에 코로나 발생으로 코로나 관련 데이터의 수가 매우 증가함을
알 수 있음

채널 별 키워드 확인

공통적으로 등장한 마스크 연관어 비교



2019년 2020년 마스크 연관어 비교

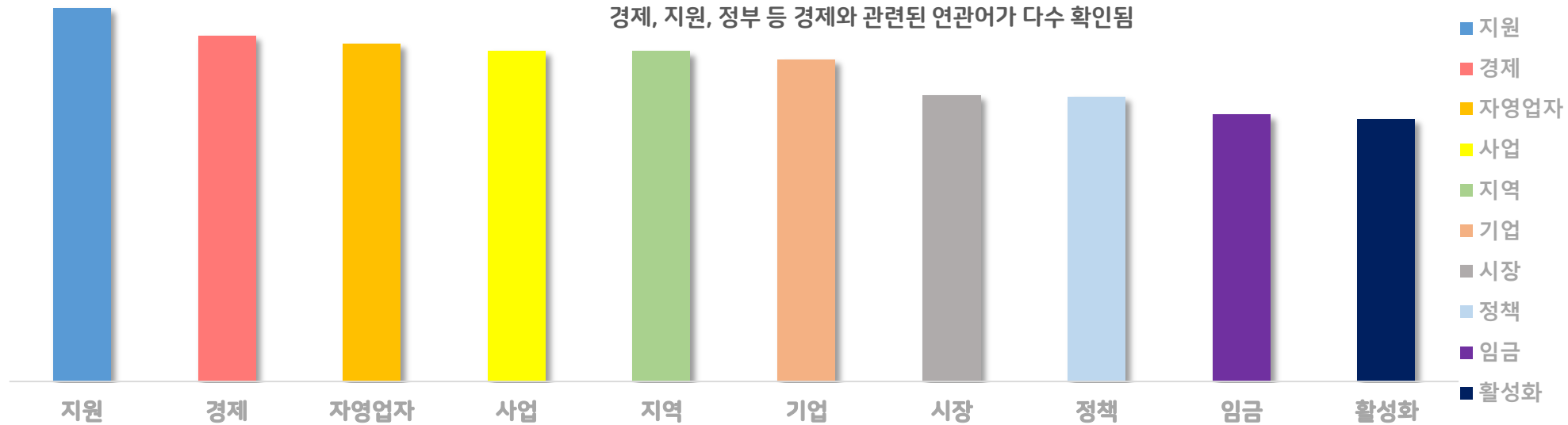
2019년에는 미세먼지, 피부, 얼굴에 관련한 단어들이 많았고 2020년에는 방역, 코로나와 관련된 단어가 많이 등장한 것을 알 수 있음

2019연관 키워드 확인

2019년의 '마스크', '소상공인', '피시방' 중 **소상공인**

소상공인과 관련된 연관어

경제, 지원, 정부 등 경제와 관련된 연관어가 다수 확인됨

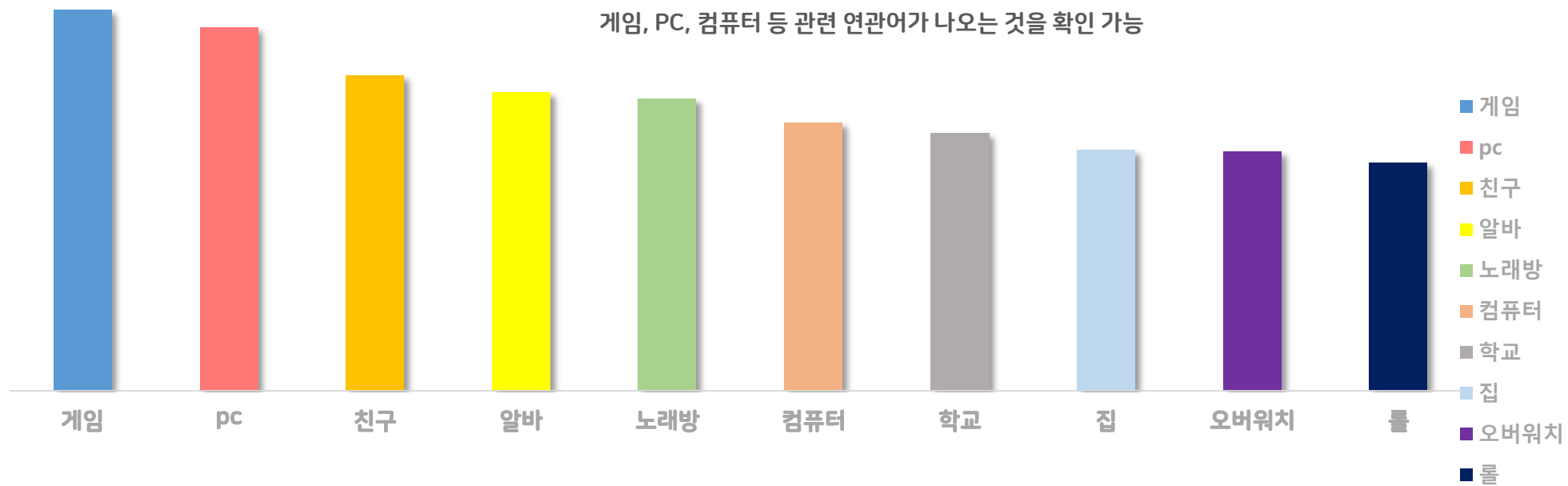


2019연관 키워드 확인

2019년의 '마스크', '소상공인', '피시방' 중 피시방

피시방과 관련된 연관어

게임, PC, 컴퓨터 등 관련 연관어가 나오는 것을 확인 가능

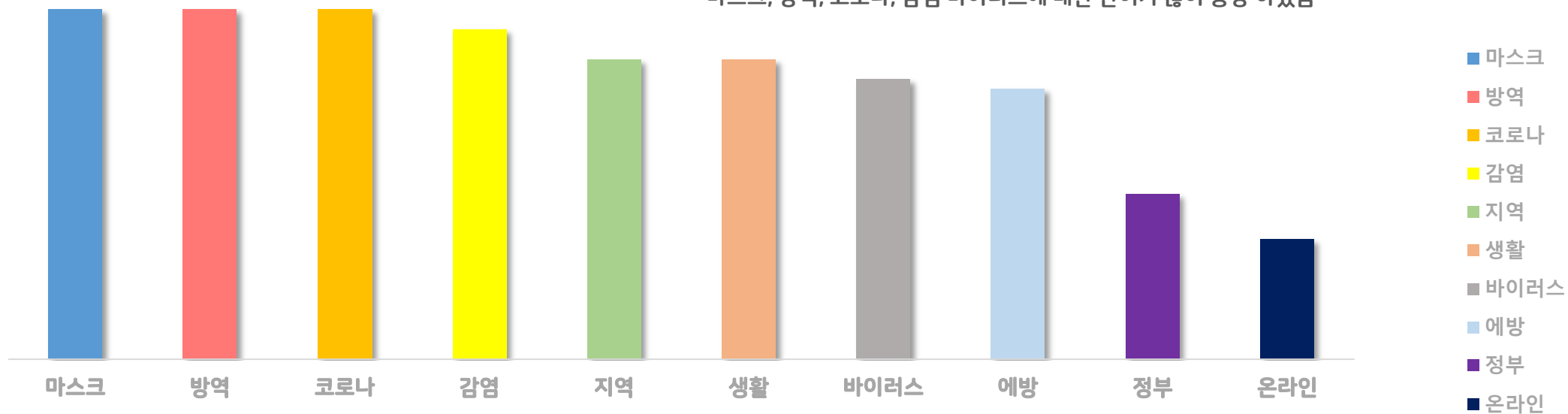


2020연관 키워드 확인

2020년의 '마스크', '코로나19', '사회적거리두기' 중 **코로나19**

코로나19와 관련된 연관어

마스크, 방역, 코로나, 감염 바이러스에 대한 단어가 많이 등장 하였음

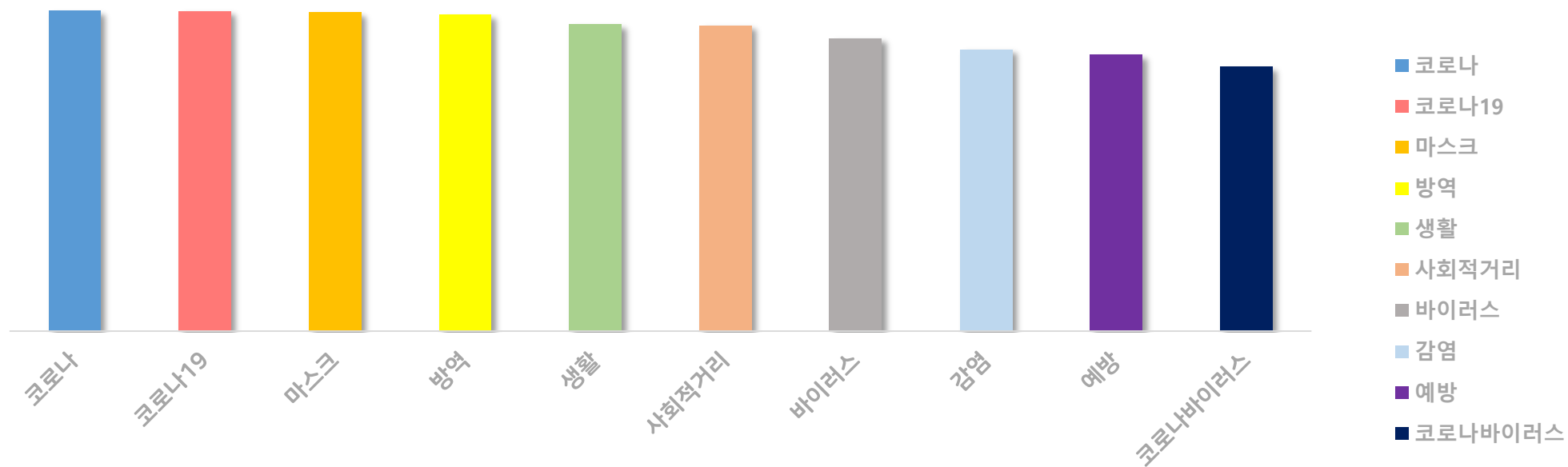


2020연관 키워드 확인

2020년의 '마스크', '코로나19', '사회적거리두기' 중 사회적거리두기

사회적거리두기와 관련된 연관어

코로나, 마스크, 방역, 바이러스, 감염이 많이 등장



2019키워드와 2020키워드

2019년의 키워드와 2020년의 키워드

- 2019년에는 마스크, 소상공인, PC방이 주요 키워드였고 2020년은 코로나, 사회적 거리두기, 마스크가 주요 키워드 였다.
- 같은 키워드임에도 불구하고 2019년의 마스크에서는 감염보다는 미세먼지와, 피부관리에 대한 연관어가 많이 등장하였고 2020년에는 코로나 바이러스에 대한 연관어가 많이 등장하였다.
- 따라서 2020년에 코로나의 영향을 많이 받고 있음을 알 수 있다.
- 다음으로는 2019년의 키워드와 2020년의 키워드를 이용하여 주식종목을 분석하였다.

• 소비데이터로 보는 주식종목 분석 •

1

종목 정리

소비데이터와 연관된 주식종목을 추리기 위해서
소비데이터의 우편번호와 주식종목의 우편번호가 겹치는 종목으로 선정

2

이슈 내용 확인

해당 주식 종목들의 이슈 내용을 먼저 확인

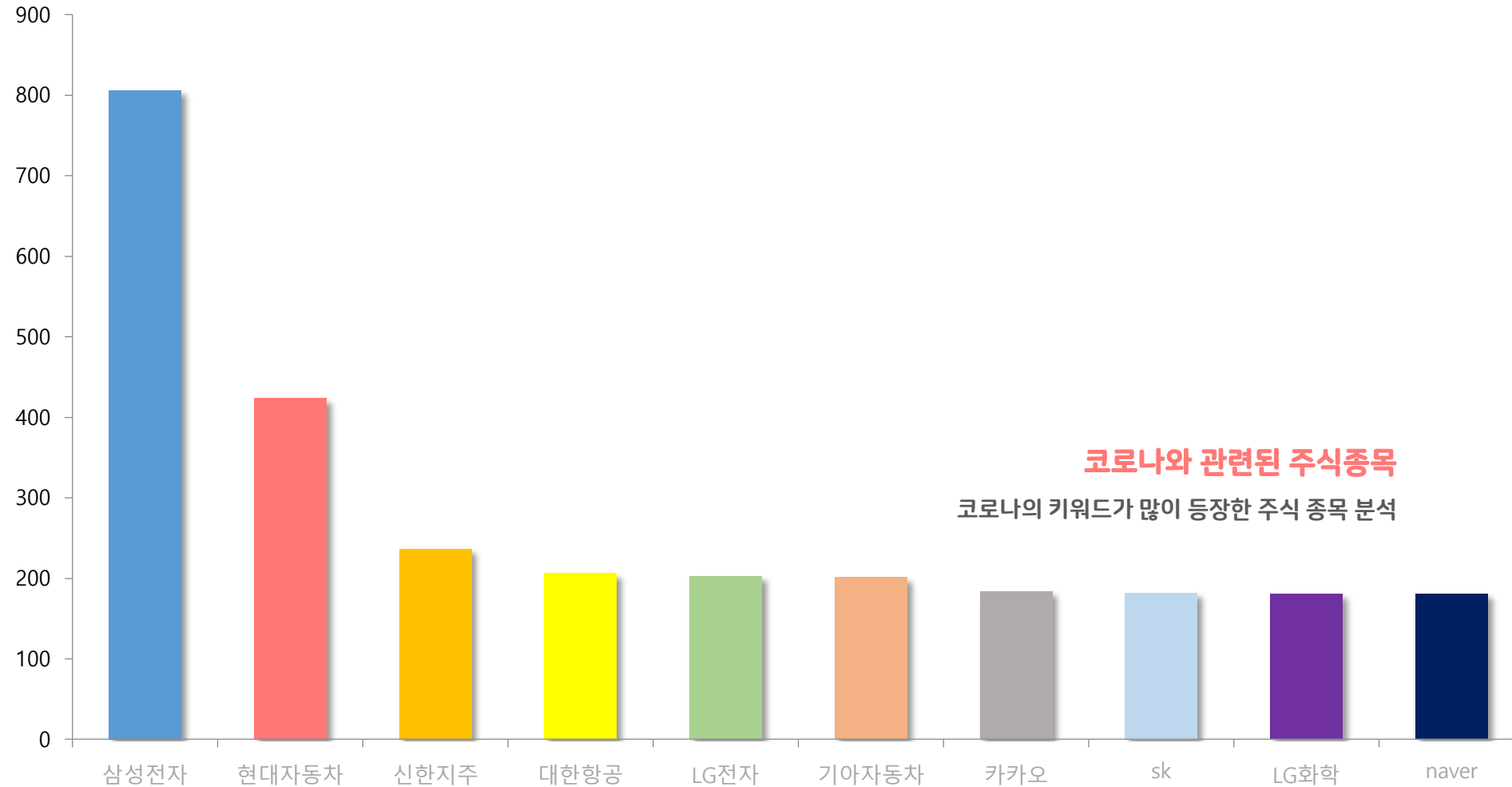
3

키워드 정리

이슈 내용의 키워드를 정리하여 그래프 분석

주식종목 분석

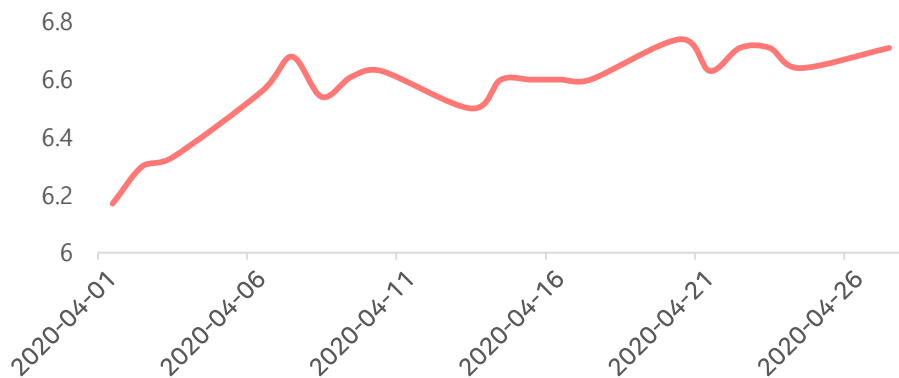
앞의 분석을 통해 코로나와 관련된 주식종목 분석



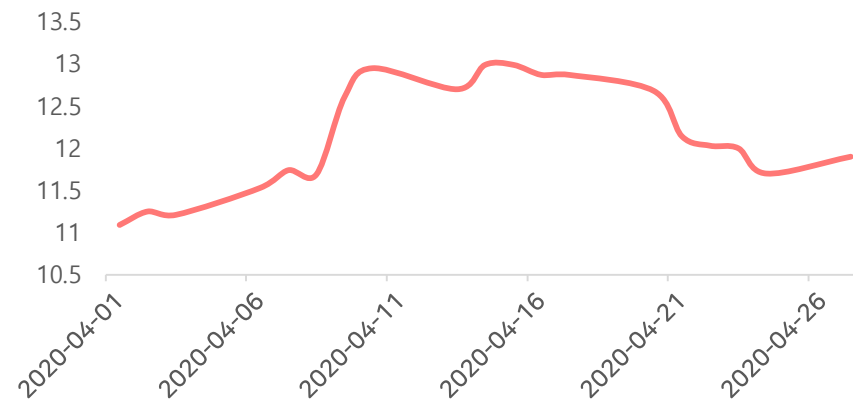
주식종목 분석

코로나를 포함한 이슈분류 문서수의 상위10종목 수익비율 변화

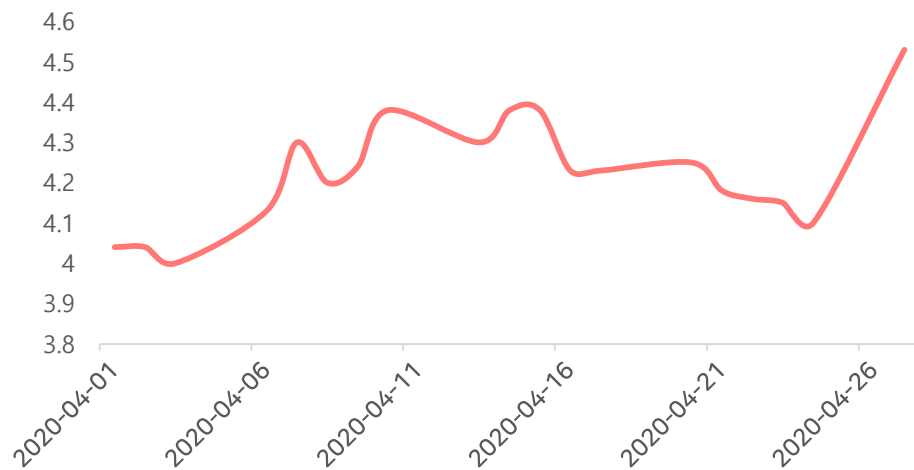
삼성전자



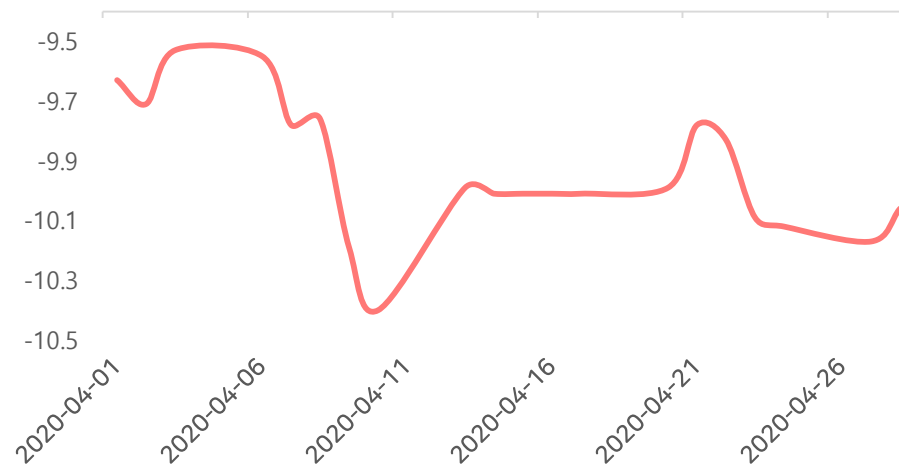
현대차



신한지주



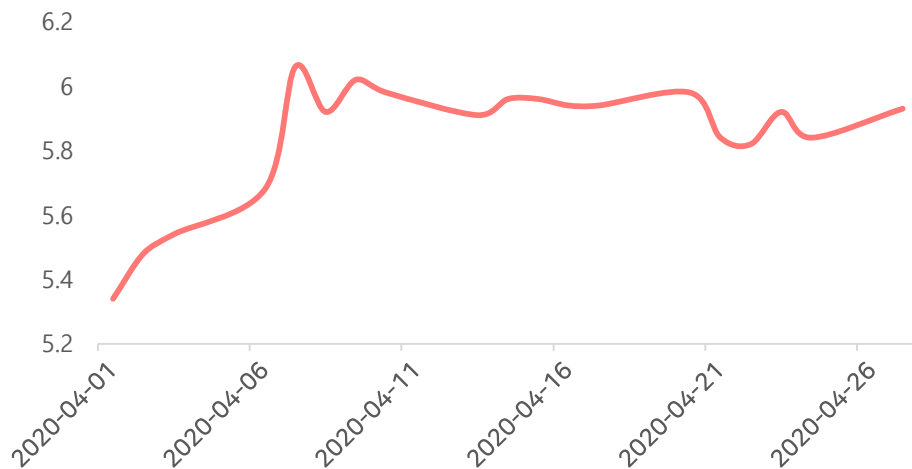
대한항공



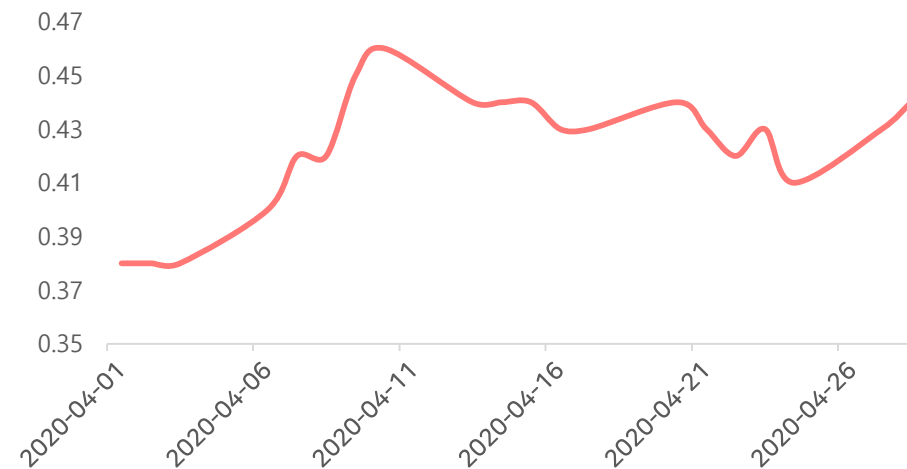
주식종목 분석

LG전자

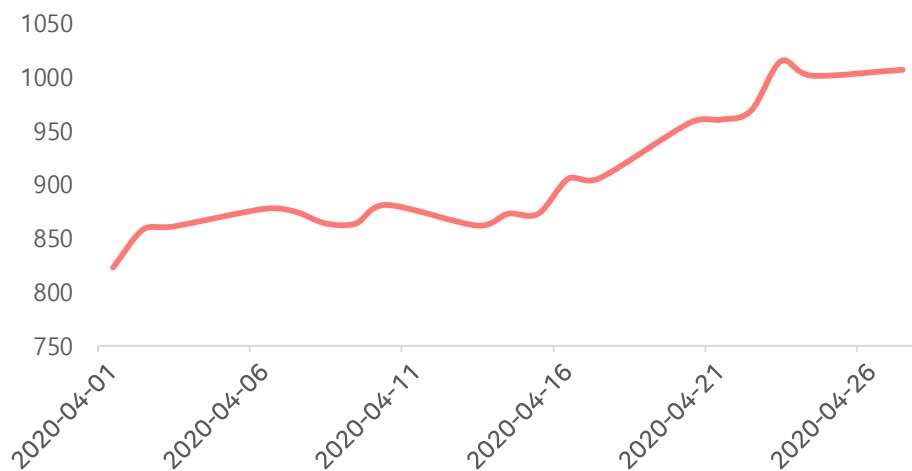
코로나를 포함한 이슈분류 문서수의 상위10종목 수익비율 변화



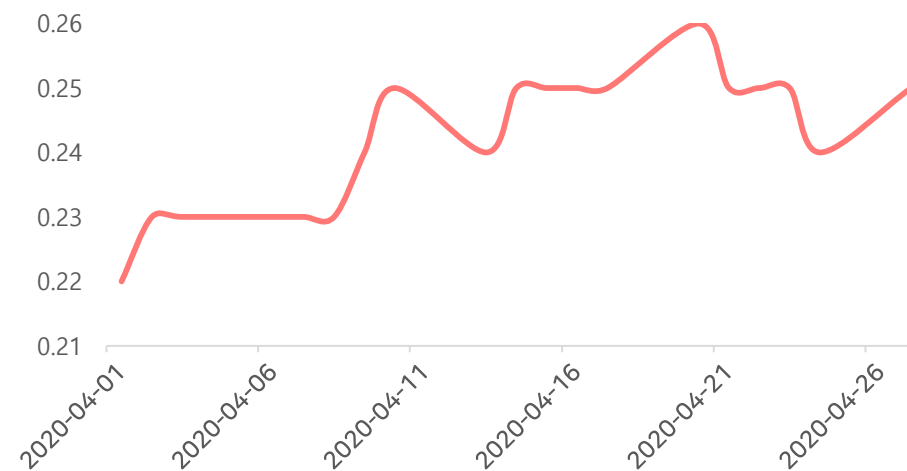
기아차



카카오



sk

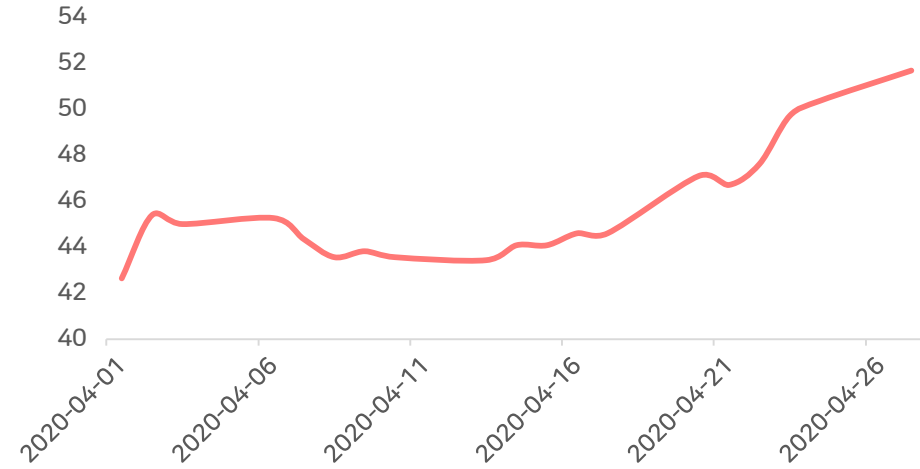
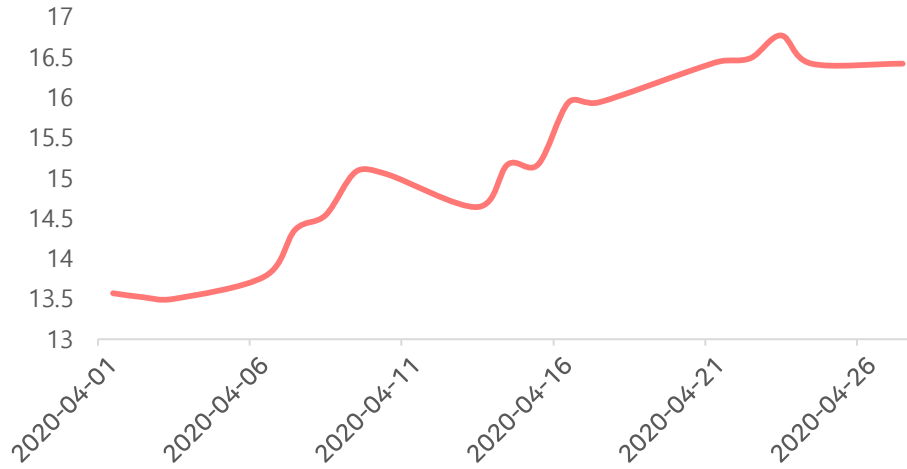


주식종목 분석

LG화학

코로나를 포함한 이슈분류 문서수의 상위10종목 수익비율 변화

NAVER



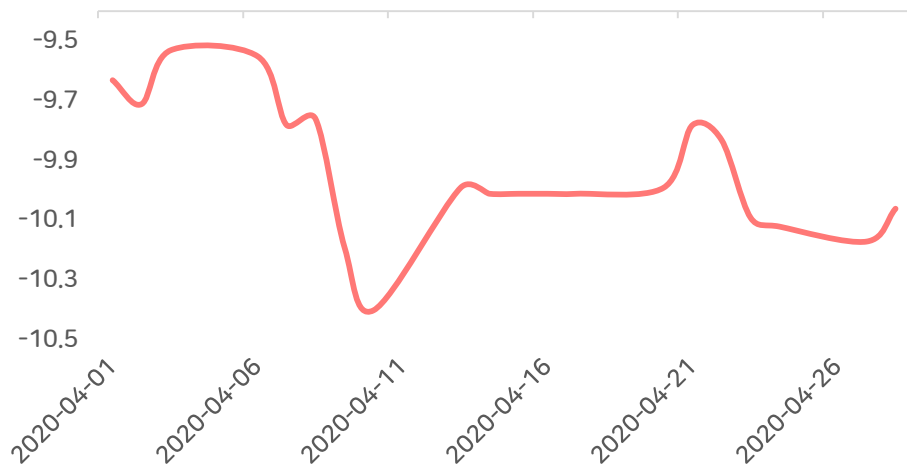
코로나를 포함한 이슈분류이용문서수의 상위10종목의 수익비율 변화 차트

주로 화학, IT, 전자 관련 주식종목들의 상승을 확인할 수 있음

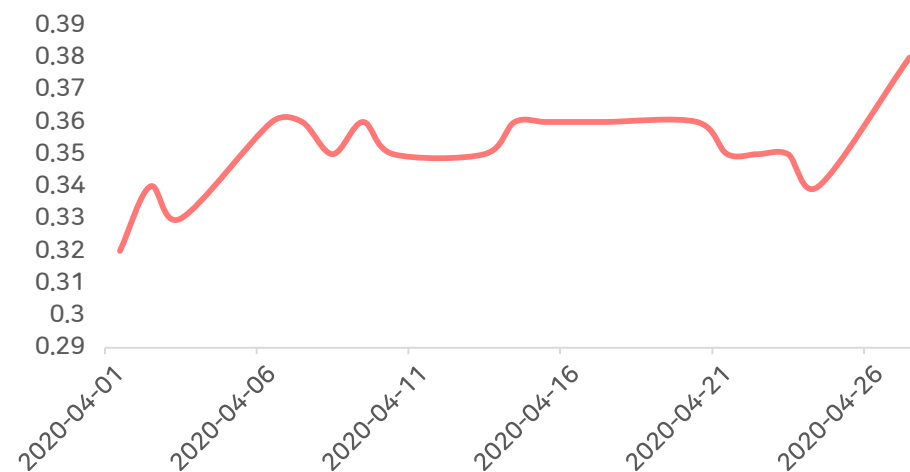
주식종목 분석

마스크를 포함한 이슈분류 문서수의 상위10종목 수익비율 변화

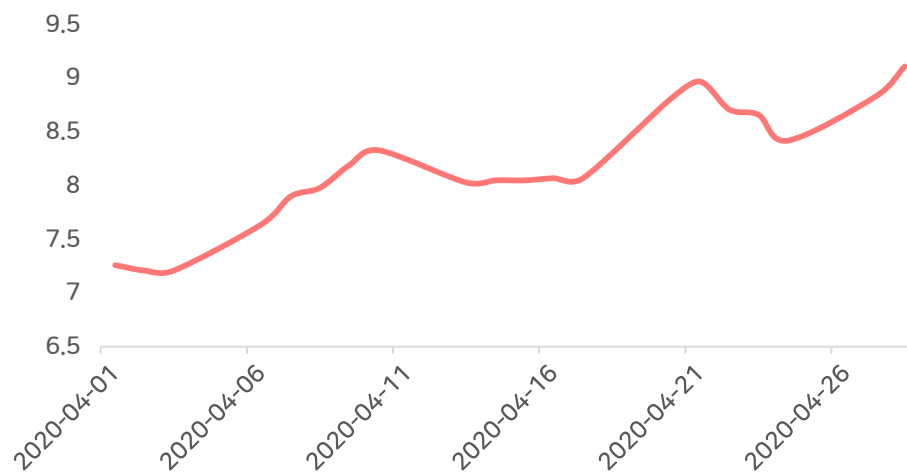
대한항공



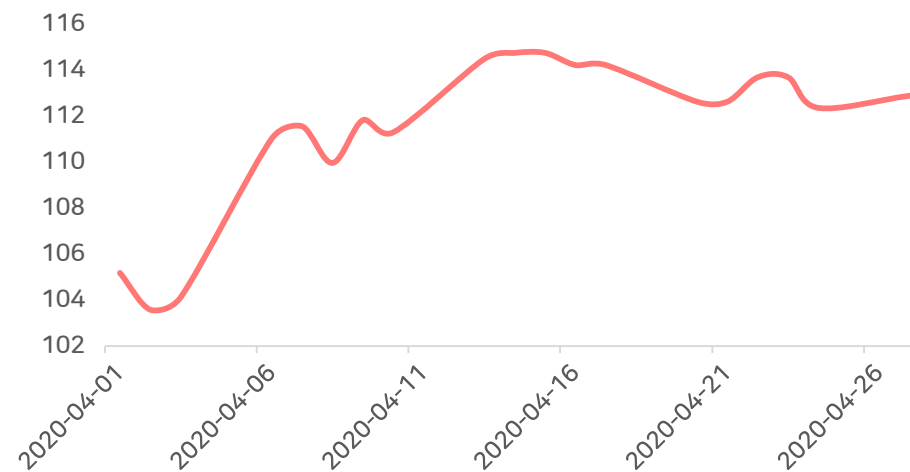
이마트



신세계



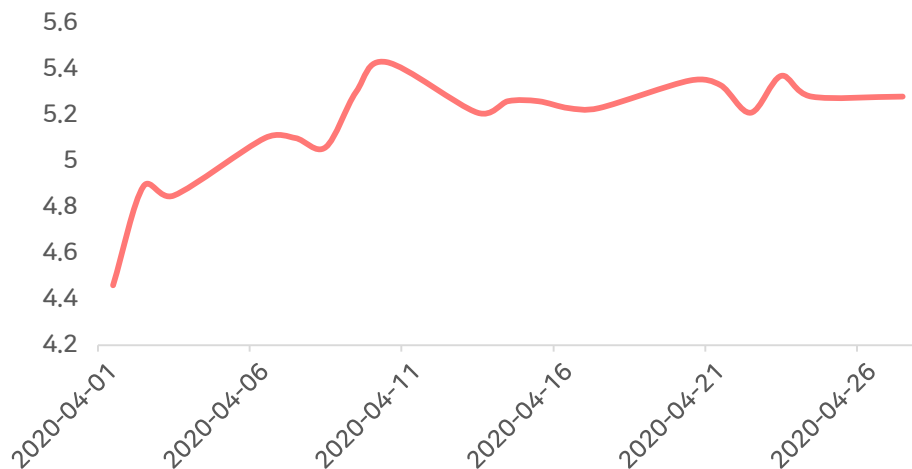
셀트리온



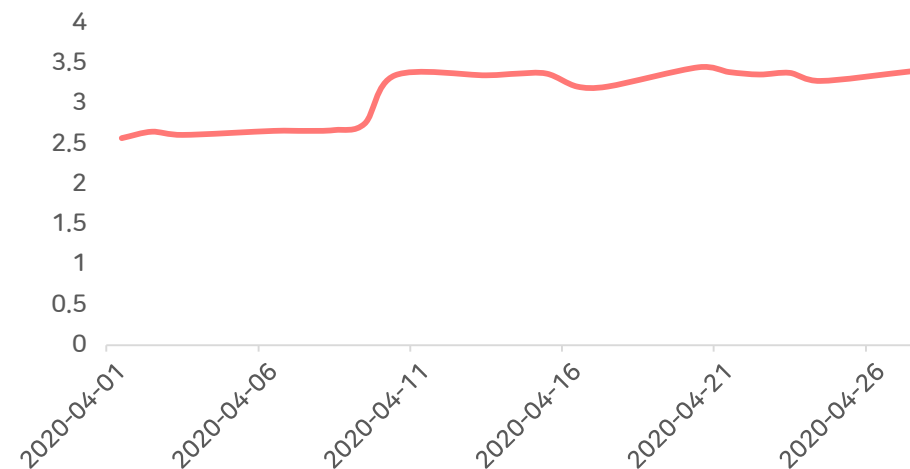
주식종목 분석

마스크를 포함한 이슈분류 문서수의 상위10종목 수익비율 변화

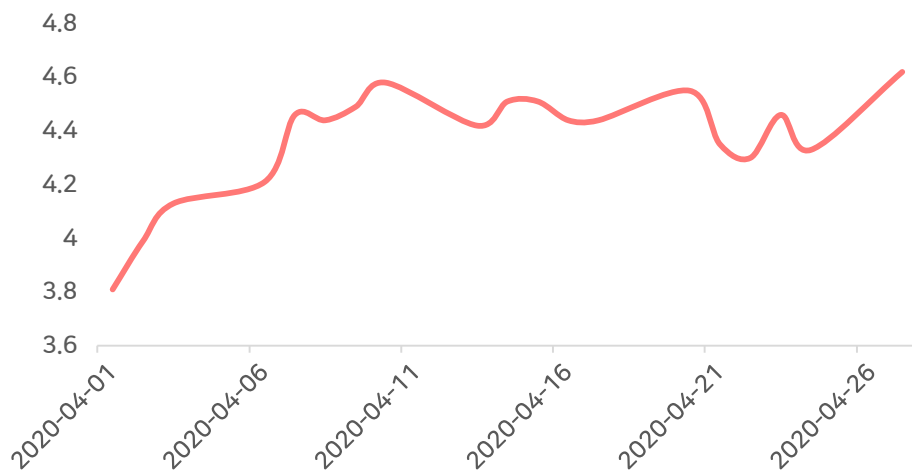
sk이노베이션



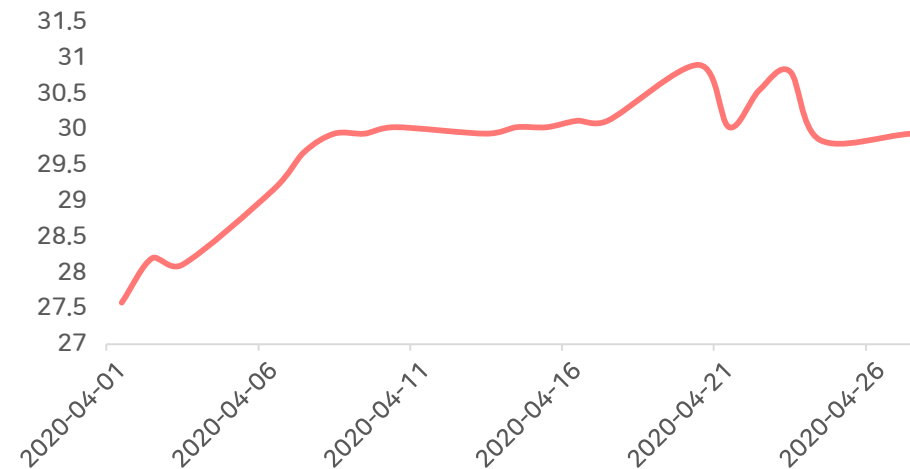
한화생명



LF



아모레퍼시픽

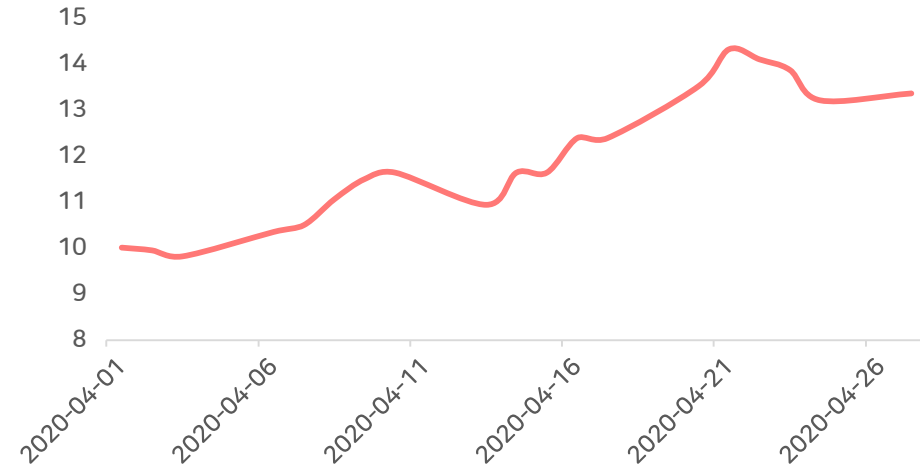
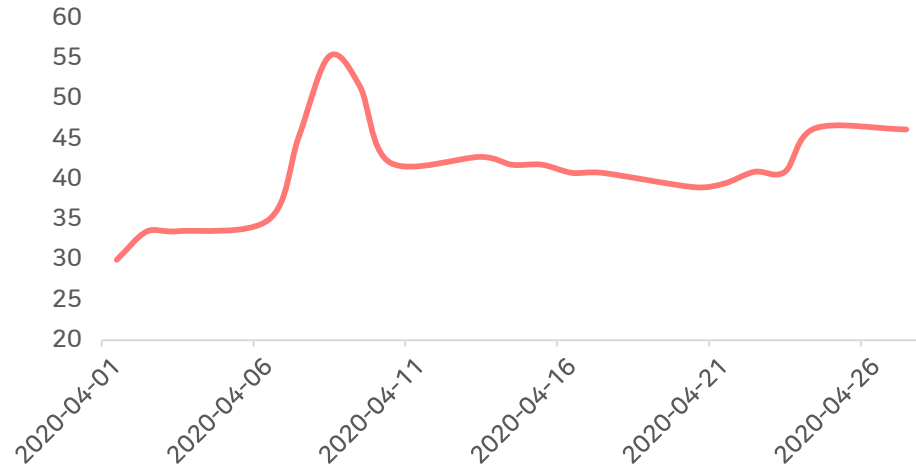


주식종목 분석

에스맥

마스크를 포함한 이슈분류 문서수의 상위10종목 수익비율 변화

skc



마스크를 포함한 이슈분류이용문서수의 상위10종목의 수익비율 변화 차트

마스크를 관련하여 판매업종, 또는 화장품, 제약 관련 종목이 상승하는 것을 확인할 수 있음

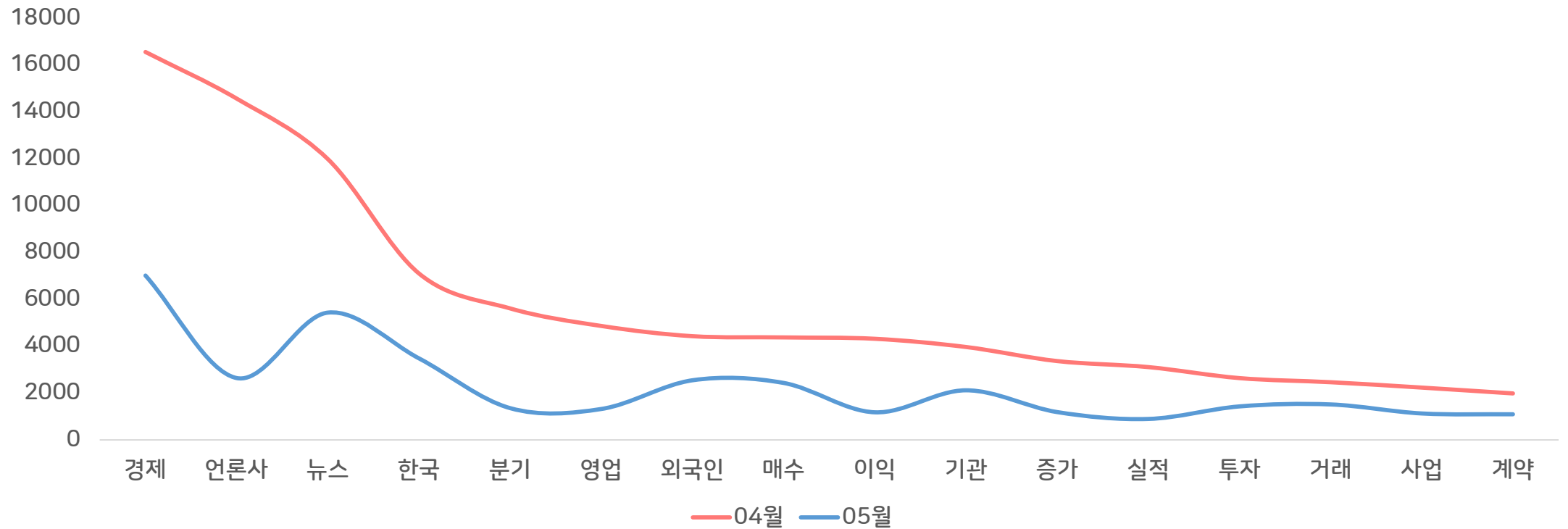
주식 키워드 분석

주식이슈데이터를 이용한 주식 키워드 분석



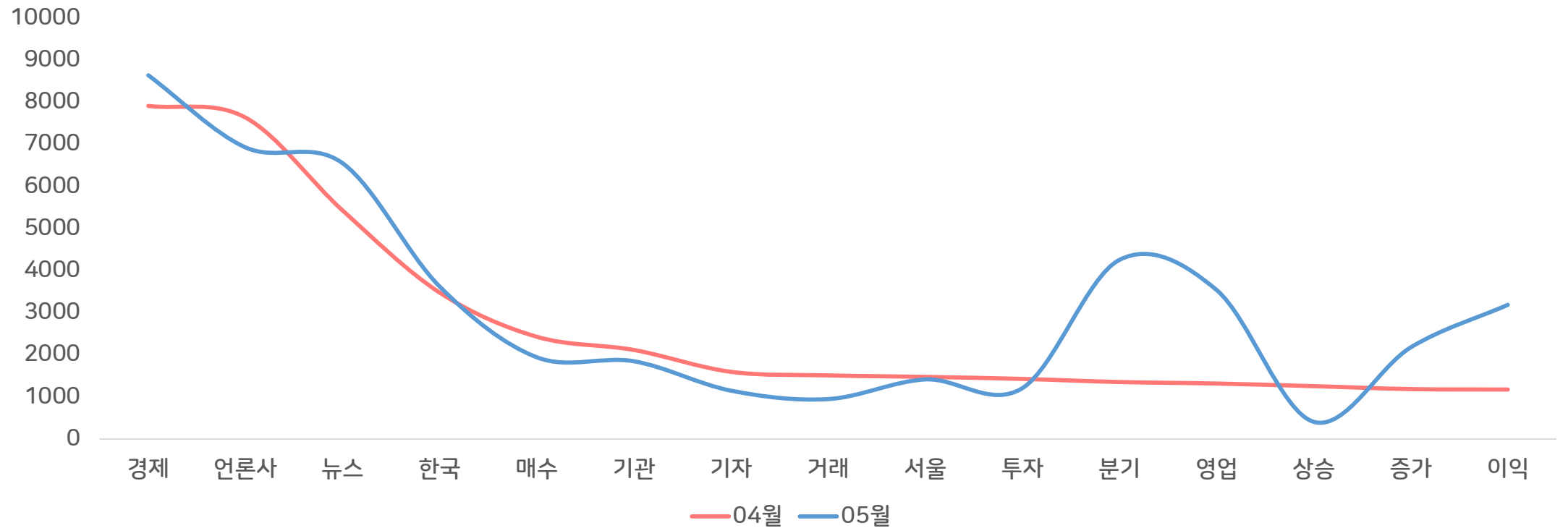
주식 키워드 분석

2019년 주식이슈데이터를 이용한 주식 키워드 분석



주식 키워드 분석

2020년 주식이슈데이터를 이용한 주식 키워드 분석



주식 키워드 분석

2019년 주식이슈데이터를 이용한 주식 키워드 분석

마스크와 연관 깊은 단어

단어	연관도
살균	0.83
셀룰로오스	0.81
방법	0.81
팩	0.79
믹스	0.77
스메	0.77
엑소	0.76
커버	0.72
쿠팡	0.71
파운데이션	0.71

경제와 연관 깊은 단어

단어	연관도
구독	0.61
뉴스	0.60
박형수	0.57
스피	0.57
이화	0.57
건조기	0.57
중앙	0.56
서연	0.56
트랙	0.55
마감	0.55

주식종목

리더스코스메틱

와이지
엔터테인먼트

락앤락

SK케미칼

프로스테믹스

애경산업

코스맥스

주식종목

삼성전자

아시아나항공

현대자동차

신한지주

기아자동차

케이티

SK텔레콤

LG전자

LG유플러스

주식 키워드 분석

2019년 주식이슈데이터를 이용한 주식 키워드 분석

미세먼지와 연관 깊은 단어

단어	연관도
오염	0.87
조작	0.82
대기	0.82
산화물	0.81
질소	0.81
배출	0.79
측정	0.79
촉매	0.77
환경부	0.75
코트	0.74

지원과 연관 깊은 단어

단어	연관도
중소기업	0.73
피해	0.72
복구	0.71
프로그램	0.70
동해시	0.70
주민	0.70
스타트업	0.69
물품	0.68
장학금	0.68
개성	0.67

주식종목

LG화학

현대제철

스튜디오드래곤

경동나비엔

한화케미칼

KC코트렐

지역난방공사

KTB투자증권

주식종목

신한지주

두산

한국전력공사

웅진코웨이

현대건설

기업은행

CJ ENM

삼성화재
해상보험

SKC

주식 키워드 분석

2020년 주식이슈데이터를 이용한 주식 키워드 분석

코로나와 연관 깊은 단어

단어	연관도
확산	0.56
정부	0.53
긴급	0.52
투데이	0.51
트레이딩	0.51
미주	0.50
감시	0.50
수출	0.50
국내	0.50
이태원	0.50

방역과 연관 깊은 단어

단어	연관도
미얀마	0.84
분사	0.84
국가	0.81
테미	0.81
전남	0.79
방식	0.78
어려움	0.76
칭호	0.76
극복	0.76
동행	0.76

주식종목

대한항공

케이티

포스코

하이트진로

한국타이어앤
테크놀로지

농심

한화

강원랜드

파미셀

주식종목

신한지주

SK

오리온

NH투자증권

한국조선해양

코오롱

SK이노베이션

강원랜드

에스엠

주식 키워드 분석

2020년 주식이슈데이터를 이용한 주식 키워드 분석

바이러스와 연관 깊은 단어

단어	연관도
균주	0.88
사이토카인	0.87
염증	0.87
효소	0.86
억제	0.86
프로테아제	0.84
열병	0.84
실험	0.84
급성	0.84

마스크와 연관 깊은 단어

단어	연관도
메타	0.70
레몬	0.67
해체	0.65
인더스	0.64
랩스	0.61
산소	0.61
방역	0.60
오펜	0.60
나노	0.59

주식종목

지트리비엔티

엠플론

젬백스

프로스테믹스

강스템바이오텍

크리스탈

코미팜

주식종목

하이트진로

한샘

LG이노텍

한국타이어앤
테크놀로지

강원랜드

SK케미칼

쌍방울

CJ프레시웨이

컴투스

주식 키워드 분석

2020년 주식이슈데이터를 이용한 주식 키워드 분석

감염과 연관 깊은 단어

단어	연관도
검체	0.93
체외	0.91
격리	0.91
일룸	0.91
정확도	0.90
피하	0.90
리브	0.90
멸균	0.90
액체	0.89

온라인과 연관 깊은 단어

단어	연관도
개학	0.76
수강	0.73
어학	0.71
와이즈	0.70
청담러닝	0.70
캠프	0.69
수업	0.69
개강	0.69
마켓	0.68

주식종목

인트로메틱

시너지
이노베이션

마크로젠

엔바이오니아

우정바이오

주식종목

웅진씽크빅

YBM넷

비상교육

알서포트

대교

손오공

다나와

인프라웨어

한국팩키지

주식 키워드 분석

2020년 주식이슈데이터를 이용한 주식 키워드 분석

지역과 연관 깊은 단어

단어	연관도
난방	0.80
어려움	0.75
집단	0.74
취약	0.72
계축	0.71
성금	0.71
경북	0.71
대구	0.70
냉방	0.69
편식	0.69

정부와 연관 깊은 단어

단어	연관도
법	0.73
방역	0.71
상황	0.71
조치	0.67
유예	0.67
국가	0.66
도입	0.66
마힌드라	0.66
휴업	0.65
분사	0.65

주식종목

GS건설

우리금융지주

현대건설

LG화학

GS리테일

한국가스공사

CJ대한통운

동부건설

SK네트웍스

주식종목

삼성전자

한진칼

GS건설

카카오

오리온

우리금융지주

하노하

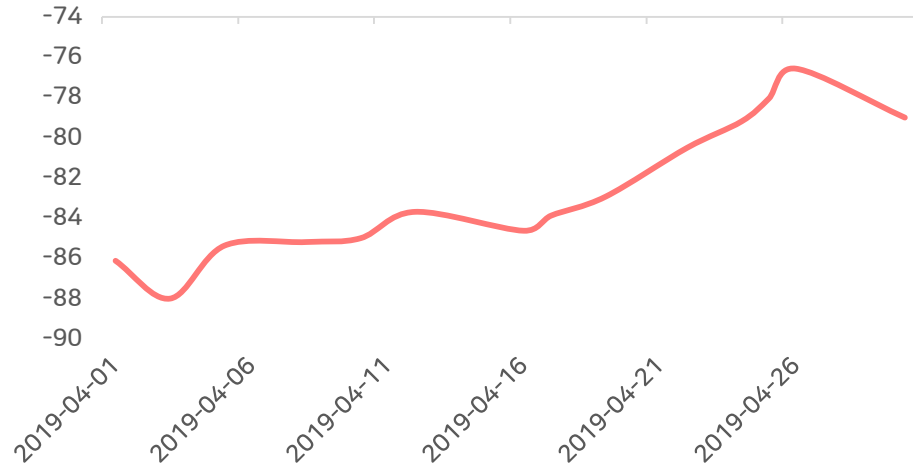
한국조선해양

셀트리온

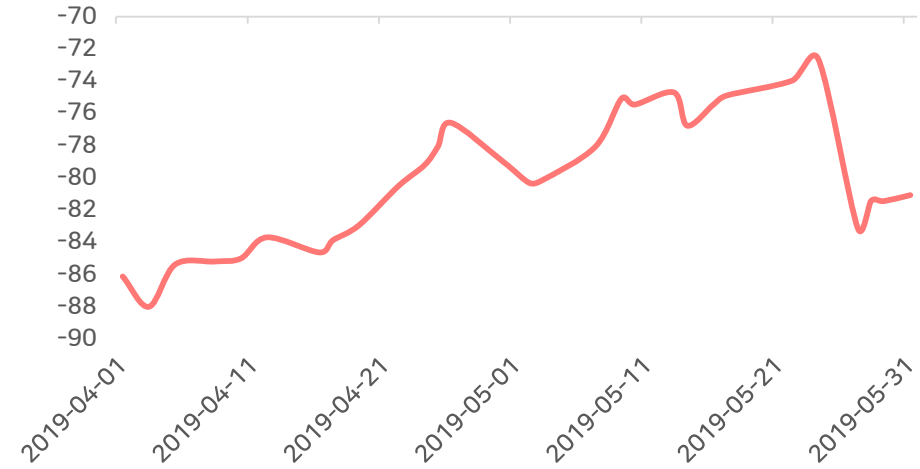
주식종목 분석

2019년 2020년의 주식 수익비율변화

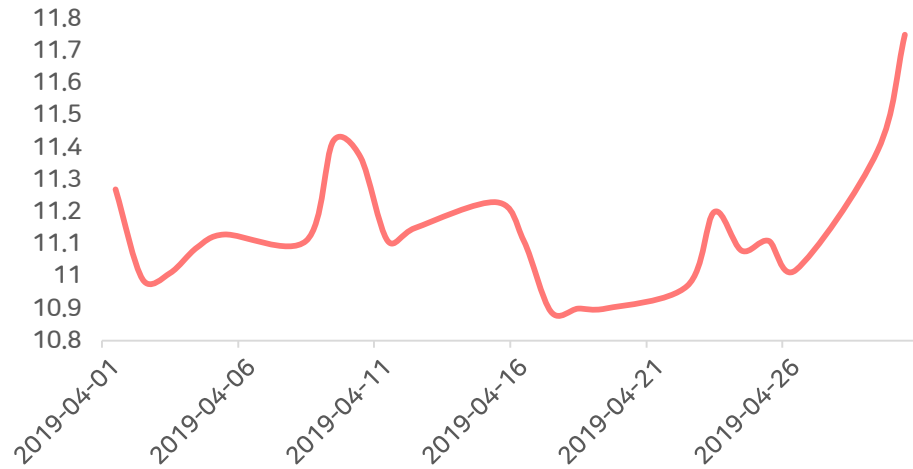
코미팜20



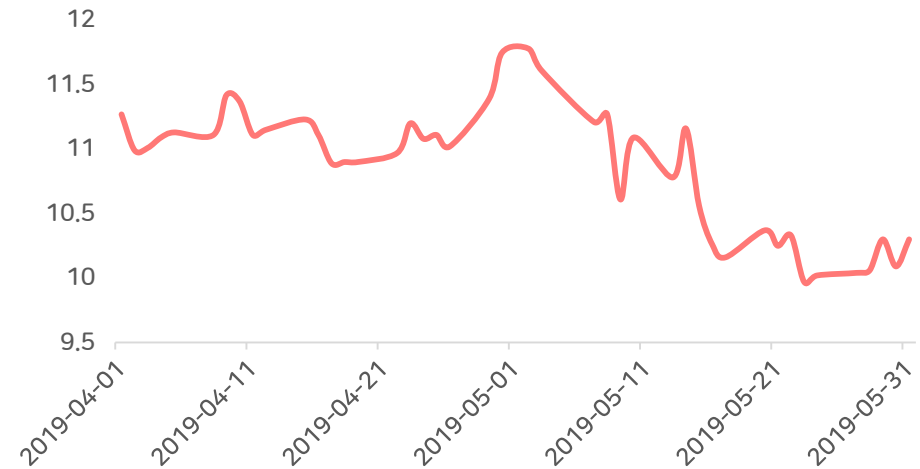
코미팜20



신세계2020



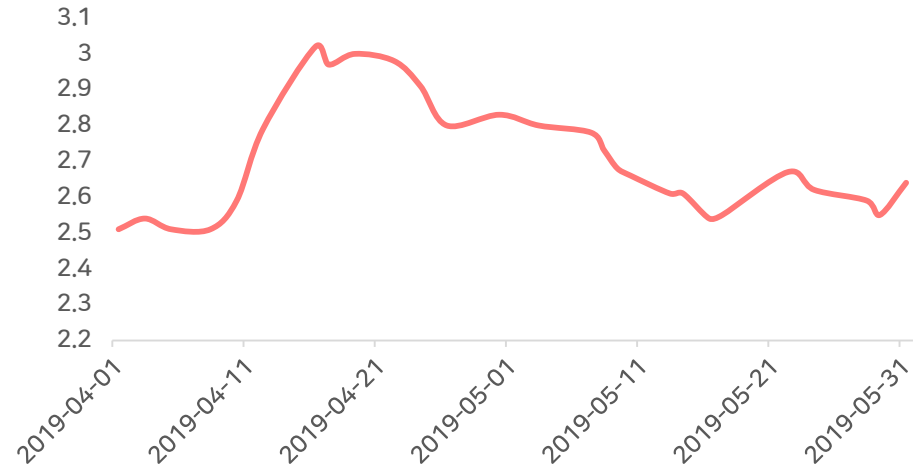
신세계19



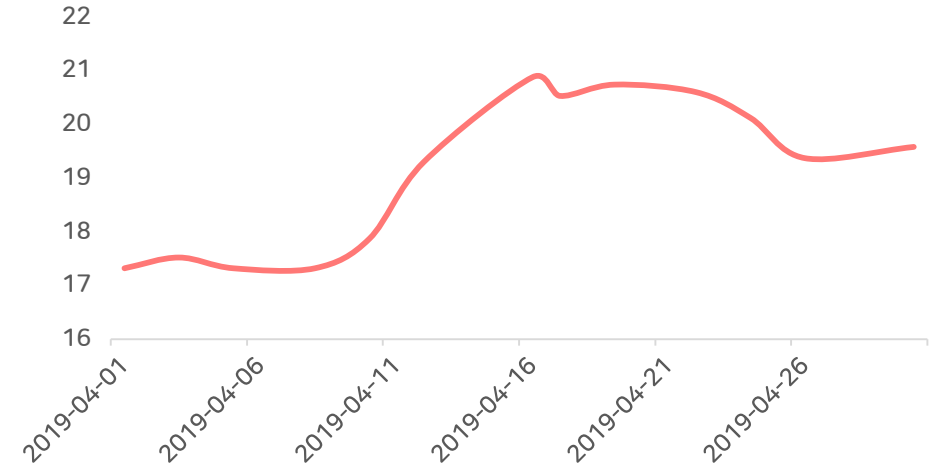
주식종목 분석

2019년 2020년의 주식 수익비율변화

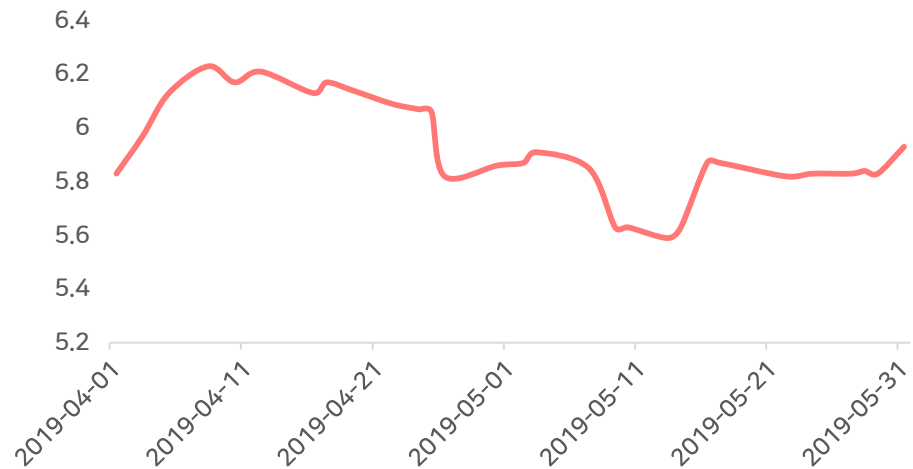
2019 nhn한국사이버



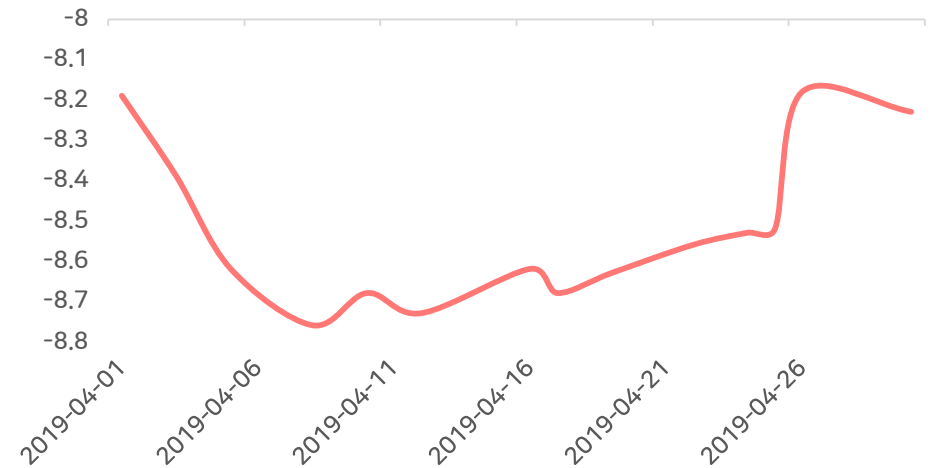
nhn한국사이버결제 20



한빛소프트2019



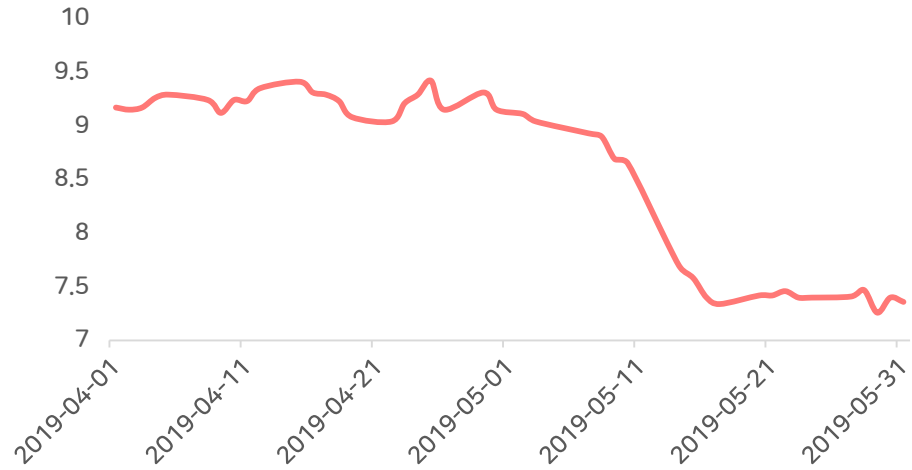
한빛소프트20



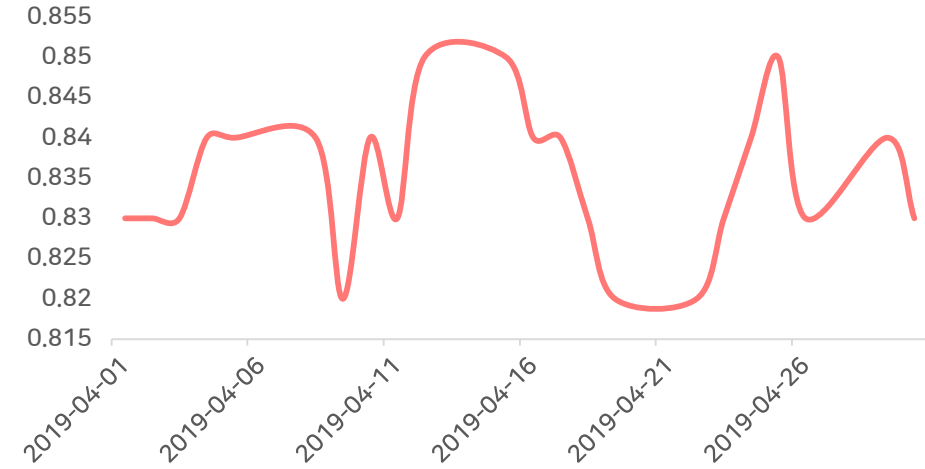
주식종목 분석

2019년 2020년의 주식 수익비율변화

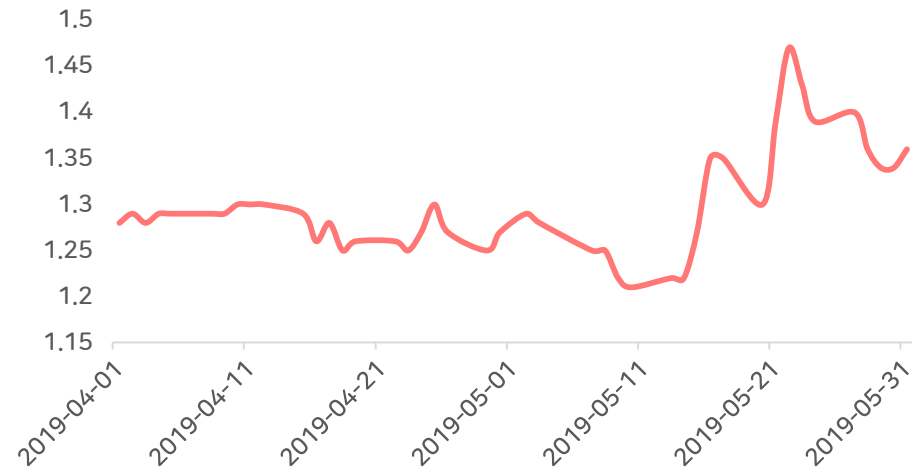
현대해상19



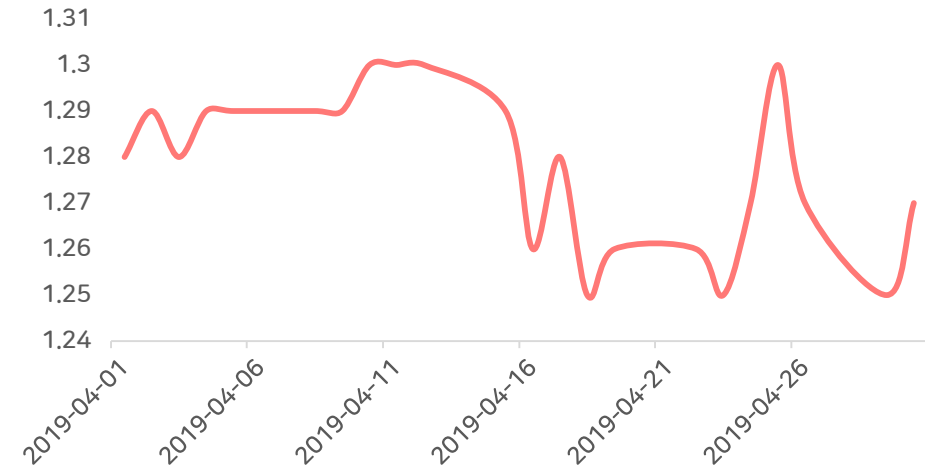
현대해상20



DB2019



db 2020



결론

주식종목 분석 결과

- 소비 데이터 분석을 통해 매출은 줄었더라도 온라인 쇼핑에 대한 거래량이 증가한 것을 알 수 있다.
- 또한 주식이슈 데이터에서도 온라인, 제약, 코로나와 관련된 단어가 등장하였다.
- 또한 코로나 바이러스로 인해 등교금지와, 비대면 상담 등 비대면 서비스가 증가하였다.
- 온라인 상거래나, 온라인 수업을 하는 교육, 온라인 거래를 위한 결제 시스템, 비대면 관련 주식 또는 IT가 많이 상승할 것이다.
- 제약과 바이오 쪽도 좋지만 코로나 사태가 나아진다면 떨어질 수 있는 가능성이 있어 유망직종에는 어려움이 있다.
- 사회적으로 좋은 일을 하거나 기부하는 경우와 크게 관련이 있어 보이지 않는다.