

데이터 사이언스 프로젝트

9조 데믹이들

강영민, 김성윤, 이성일, 이호건









Go Beyond Education

동아출판



한빛미디어 HanbitMedia, Inc.



21세기북스



키워드, 상품 코드 검색

🏠 홈 > 시장보고서 > 일반소비재

시장보고서

상품코드 1076888



세계의 출판 시장(2022-2026년)

Global Publishing Market 2022-2026

발행일: 2022년 05월 | 리서치사: > TechNavio | 페이지 정보: 영문 120 Pages | **배송안내 : 즉시배송**



자료문의 >



열람서비스 >



샘플신청 >

☐ 비교하기

※ 본 상품은 영문 자료로 한글과 영문 목차에 불일치하는 내용이 있을 경우 영문을 우선합니다. 정확한 검토를 위해 영문 목차를 참고해주시기 바랍니다.

세계의 출판 시장 규모는 2022-2026년 192억 달러의 증가가 전망되며, 예측 기간 중 CAGR로 1.22%의 성장이 예측됩니다.

시장은 전자서적의 영향 확대, 인터넷 보급과 스피드의 향상, 출판 업계의 비즈니스 역학 변화 등에 의해 촉진되고 있습니다.

[표 01] 최근 3년간(2020~2022년) 주요 출판사(77개 사)의 매출액 현황

(단위 : 백만 원, %)

2022년 출판시장 통계

77개 출판 기업 및 서점을 통해 본
2022년 출판시장 동향 분석 보고서

2023. 04.



순위		기업		매출액		이전 매출액		분류	비고
2022	2021	이름	형태	2022	증감률	2021	2020		
69	70	(주)스토리위즈	기타 법인	10,725	30.7	8,207	5,454	만화·웹툰·웹소설 출판	
70	66	(주)기탄교육	기타 법인	9,099	-16.1	10,847	15,039	교육 출판	
71	69	지앤비패럴렉스교육(주)	기타 법인	8,774	-1.2	8,883	7,097	교육 출판	구 지앤비교육(주)
72	72	(주)제이스튜디오	기타 법인	7,668	30.0	5,898	4,742	만화·웹툰·웹소설 출판	
73	74	(주)단꿈아이	기타 법인	6,232	33.1	4,683	8,441	단행본 출판	
74	73	(주)도서출판한국헤르만헤세	기타 법인	5,546	-3.8	5,764	3,961	교육 출판	
75	76	(주)심슨영어사	기타 법인	2,814	-10.6	3,149	2,160	교육 출판	3월 결산
76	75	(주)가나문화콘텐츠	기타 법인	2,594	-22.1	3,330	4,000	단행본 출판	
77	77	(주)이앤미래	기타 법인	1,568	-29.7	2,231	2,562	교육 출판	
소계(77개사)				5,108,059	2.8	4,968,484	4,439,879		
참고		한국교육방송공사(EBS)	공영방송	85,596	-6.8	91,806	90,584	교육 출판	출판사업 매출액

※ 출처 : 금융감독원 전자공시시스템(<http://dart.fss.or.kr>)에 게시된 각 기업의 감사보고서

전 국민 독서캠페인
책 읽는 대한민국

책으로 세상을 비평하다
독서신문
Since 1970

☰ 전체 기획 인터뷰 책 문화일반 기관/기업 오피니언 피플



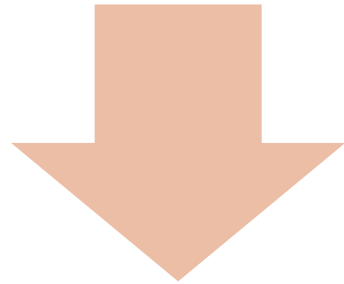
HOME > 문화일반 > 출판

출판사 영업이익 38.7% 줄었다... 출협, '2022년 출판시장 통계' 발표

김혜경 기자 | 승인 2023.04.27 10:02 | 댓글 0



성장하는 출판시장 속 영업 이익을 늘려야
한다!



지식을 다루는 출판사의 위기



“독서 트렌드와 영향 요인 파악”

<기대효과>

“독서 트렌드와 영향 요인을 파악한다면, 이를 마케팅 전략에 활용하여
영업손실에 해당하는 마케팅 비용과 인건비 등을 최소화할 수 있을 것이다.”

도서관 데이터
(도서관 정보나루)

서점 데이터
(알라딘)

뉴스 기사 데이터
(구글)

<가설>

“특정 소재가 이슈가 된다면, 이는 독서 트렌드에 영향을 미칠 것이다.”

<도서관 데이터>
대출 수
상위 1000위 데이터

<서점 데이터>
베스트 셀러
상위 1000위 데이터

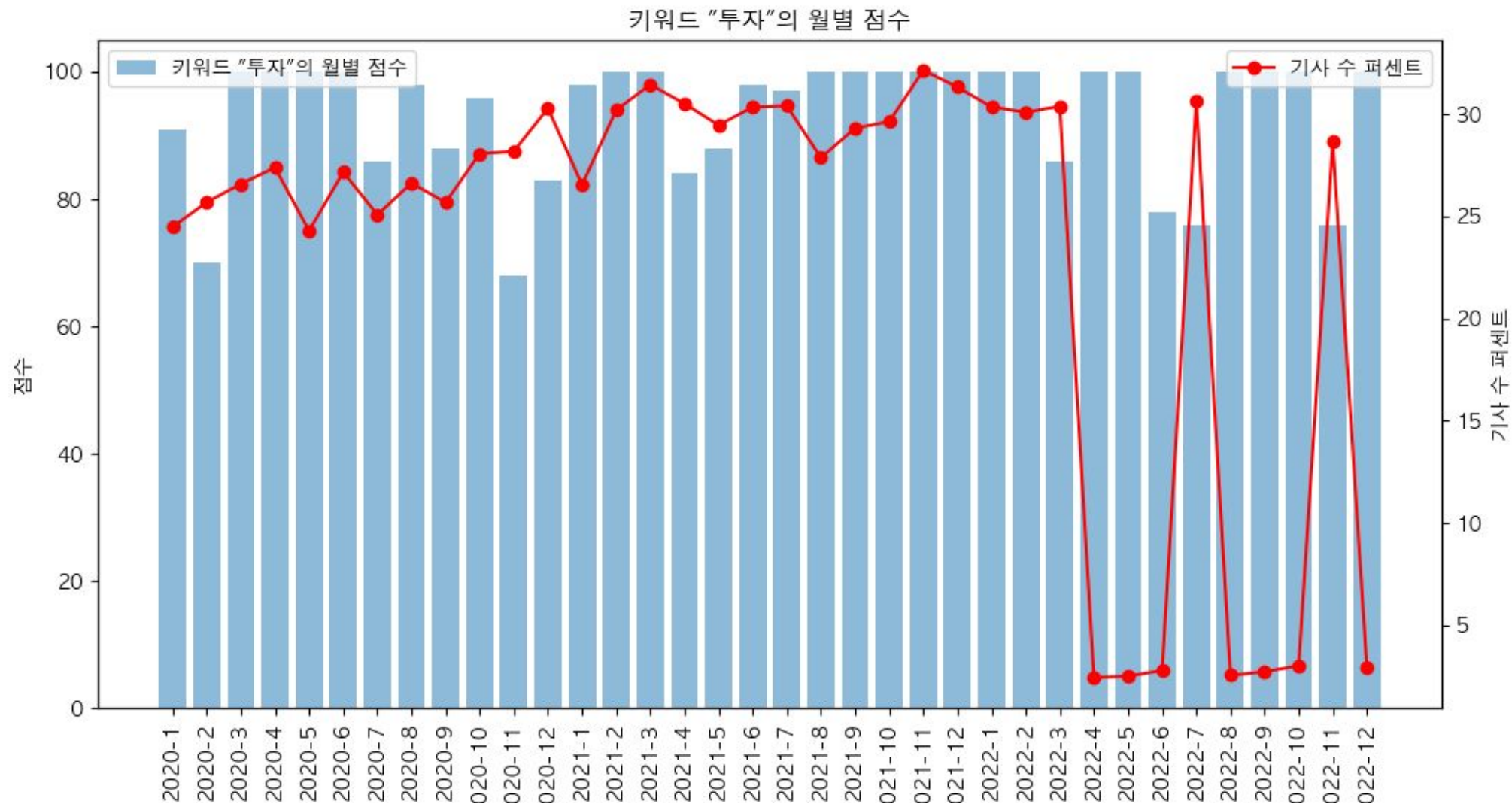
Pandas를 이용하여 전처리 및 '경제' 도서 선택

TF-IDF을 이용하여
도서 제목에서
키워드 추출

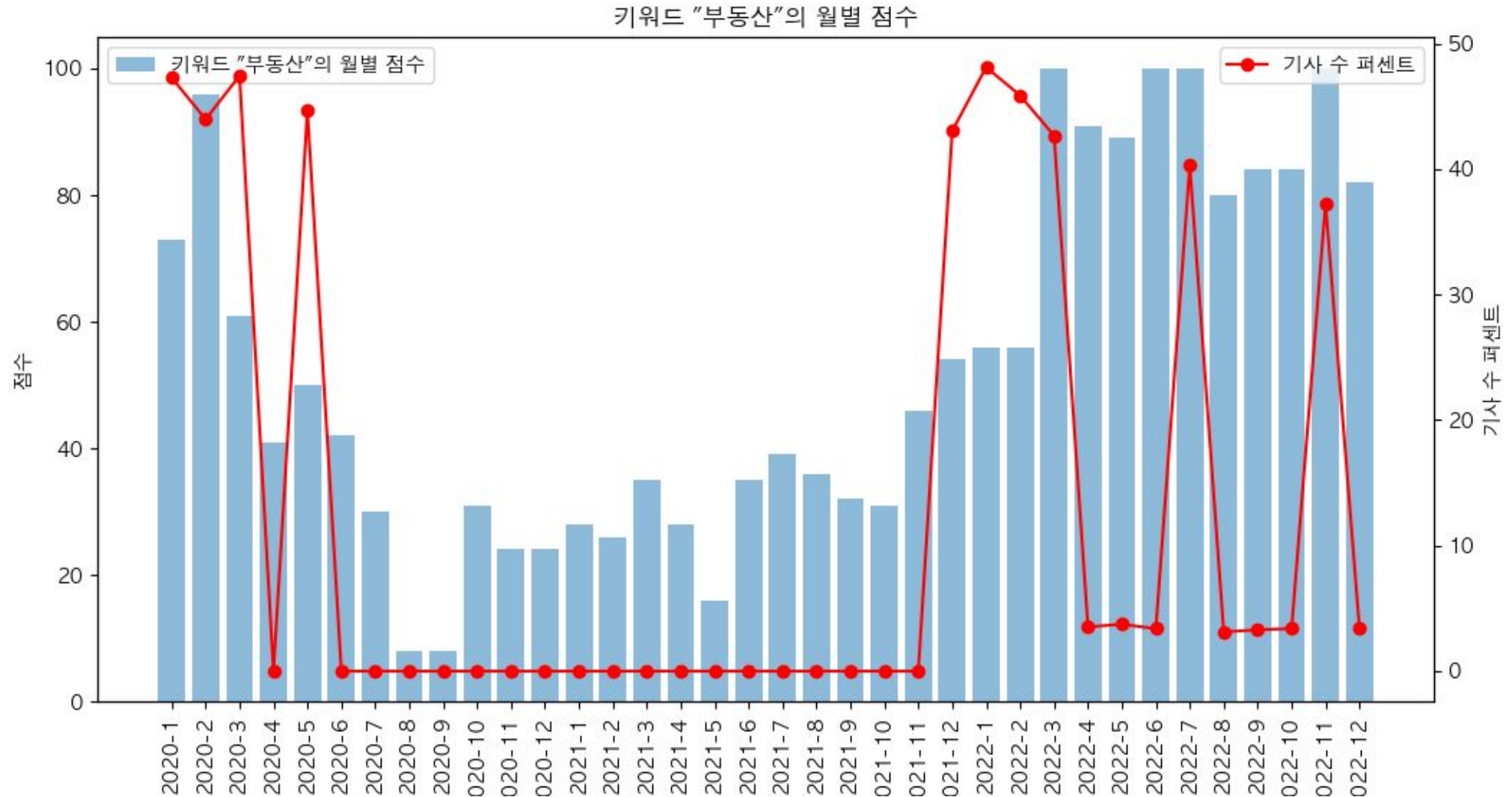
크롤링을 이용하여
키워드 별
구글 뉴스 기사 수 파악

<Loop>
2020-2022
1월-12월

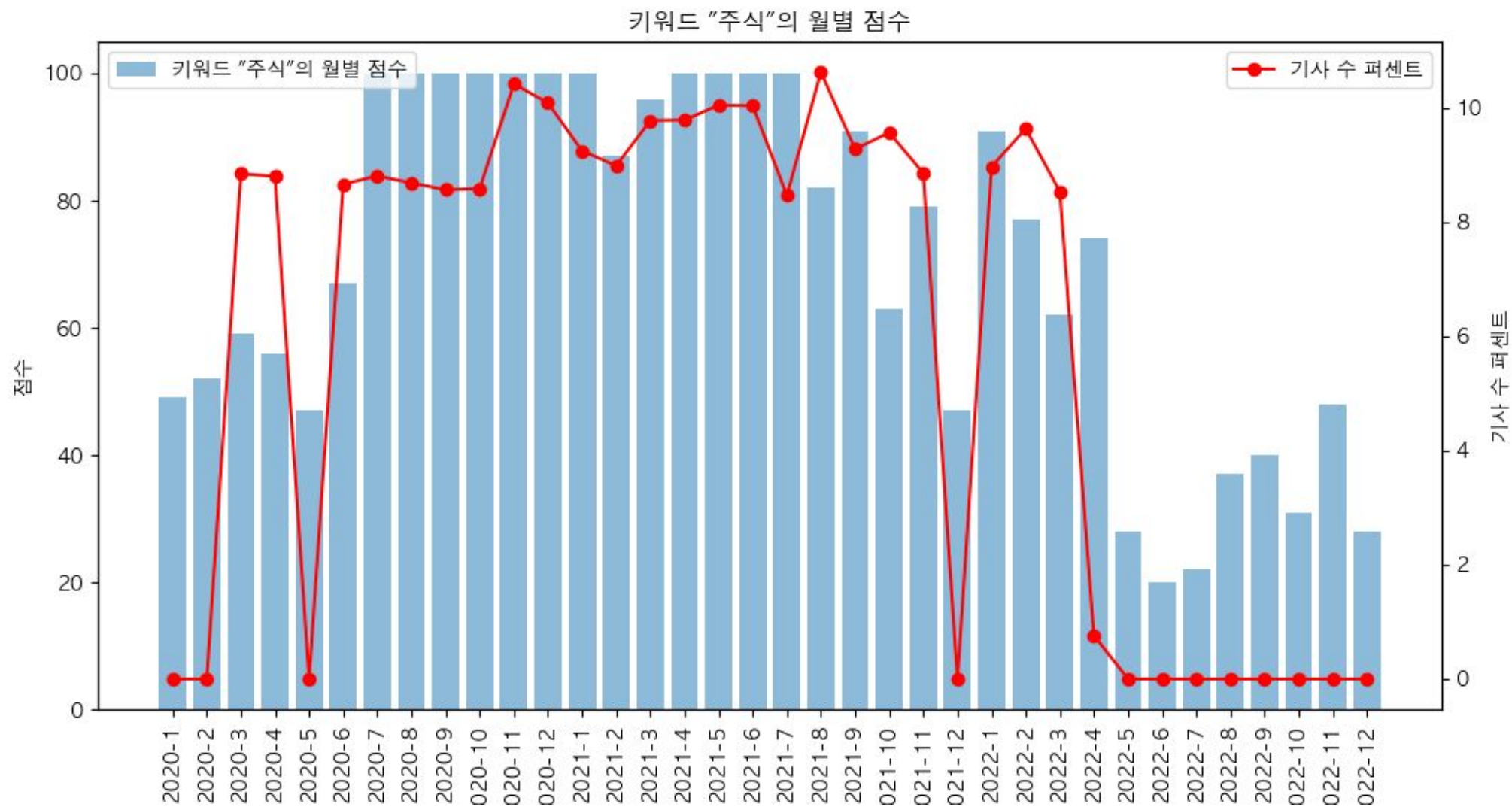
서점 데이터 - 뉴스 기사 비율



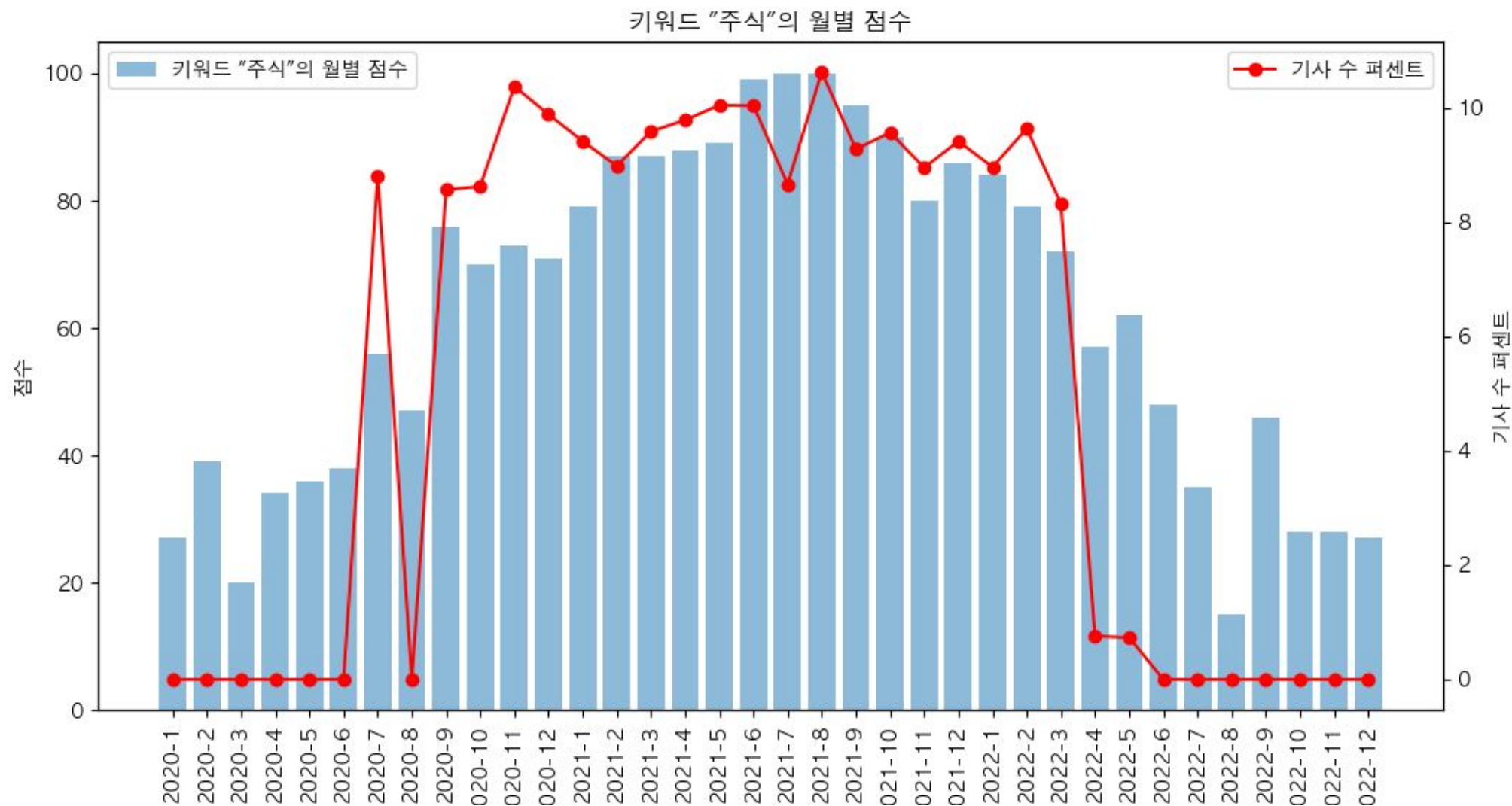
서점 데이터 - 뉴스 기사 비율



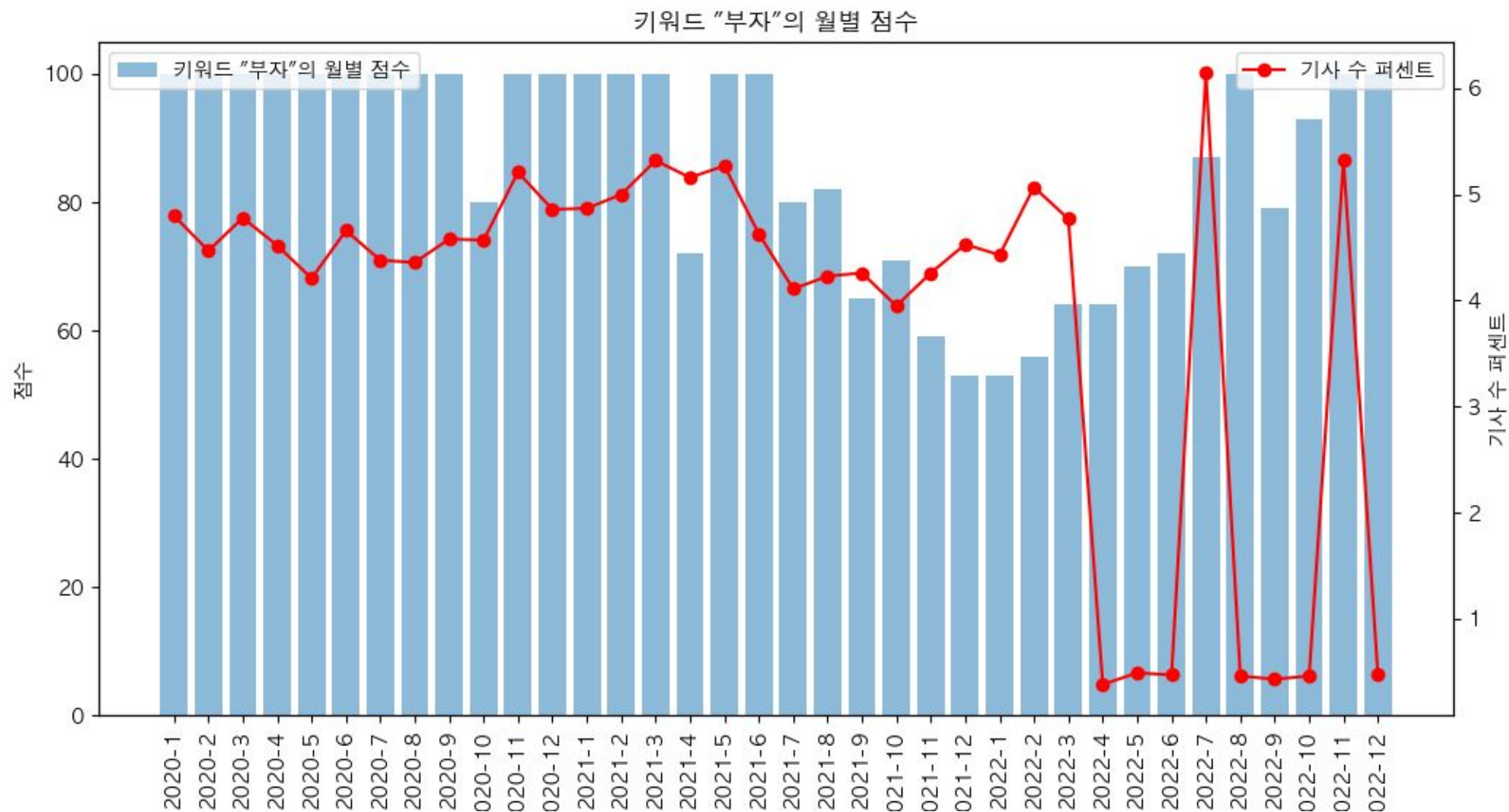
서점 데이터 - 뉴스 기사 비율



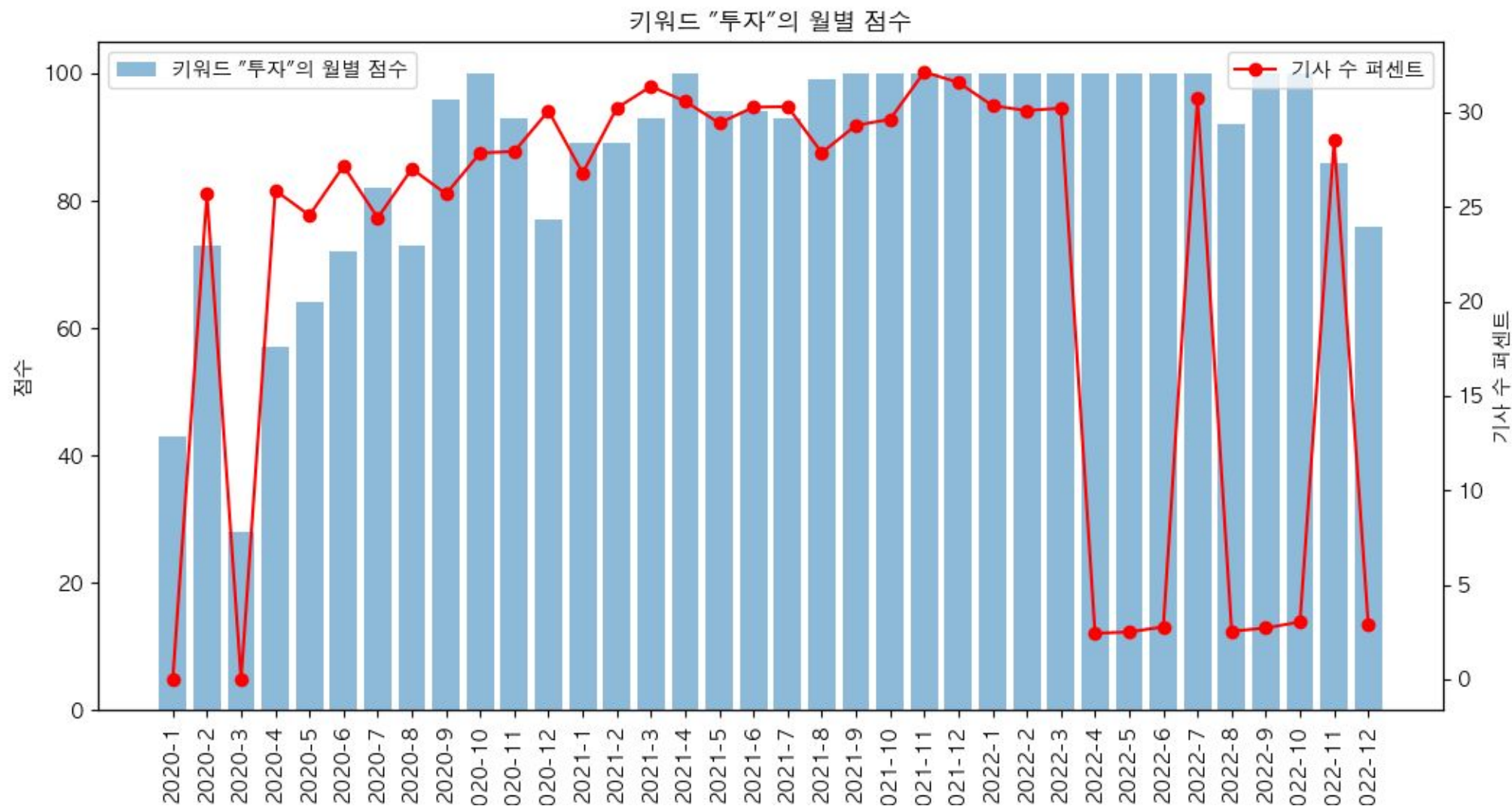
도서관 데이터 - 뉴스 기사 비율



도서관 데이터 - 뉴스 기사 비율



도서관 데이터 - 뉴스 기사 비율



<분석 결론>

“키워드마다 책의 반응 타이밍이 달랐다!”

“키워드가 확실하고 이슈가 클수록
데이터들이 밀접한 연관성을 보였다”

<가설>

“특정 소재가 이슈가 된다면, 이는 독서 트렌드에 영향을 미칠 것이다.”

⇒ 무조건적인 영향을 미치지 않고,
키워드의 특징에 따라 다양한 반응을 보였다.

[한계점]

1. 다양한 주제에서 시도하지 못했다.
2. 제목만을 활용해서 정확도가 낮을 수 있다.
3. 키워드 추출 과정에서 전처리가 완벽하지 못했다.
4. 이슈에 대한 반응이 즉각적으로 오지 않을 수 있다.

[향후 계획]

1. 경제 이외의 주제로 실험한다.
2. 도서 줄거리랑 뉴스 기사 내용까지 크롤링해서 키워드 추출에 활용한다.
3. 문제집과 같은 책은 배제하고 형태소 분석기를 활용해 전처리를 한다.
4. 새로운 가설을 제시하여 새로운 영향 요인을 파악해본다.
5. 측정 기간을 늘려서 이슈에 대한 반응이 늦게 나타나는 패턴도 파악한다.

Thank
you

부록: 팀원별 역할

<강영민>

- 네이버 기사 키워드별 크롤링 코드 작성
- 구글 기사 키워드별 크롤링 코드 작성
- 추출 키워드 데이터 분석을 통한 방향성 재모색
- 2021년 도서 데이터 주제별 추이분석
- 제안발표자, 제안발표 ppt 제작

1

.

<이성일>

- 데이터 전처리 후 분석으로 방향 수정
- 제목 자연어 분석 모듈 제시 후 통합
- 크롤링과 코드 통합

3

.

<김성윤>

- 소장, 최종발표자
- 아이디어 제안 및 방향성 제시
- 중간발표 ppt, 최종발표 ppt 제작
- 키워드 추출 함수 제작, 네이버 크롤링 코드 제작
- 핵심 데이터 파일(data.csv) 뽑아내는 코드 병합 및 제작
- 시각화 활용 파일(book_dt.csv, lib_dt.csv) 완성
- 시각화 코드 제작
- 시각화 결과 해석
- 배경, 목표, 기대효과 제시

2

.

<이호건>

- 아이디어 제안 및 프로젝트 방향성 제시
- 서점/도서관 서적 데이터 제목 자연어처리
- 20년도 데이터 담당, 중간발표자
- 데이터 시각화

4

.

