

<배경> - 강영민

<문화와 창의성의 관계> - 이호건

‘문화는 개인의 인지 과정이나 사회적 규범 및 조직을 통해 창의성에 영향을 주기 때문에(Morris & Leung, 2010) 둘은 밀접한 관계에 있다.’ - 이재윤, 김명언, 문화와 창의성 개관: 과거, 현재 그리고 미래, 한국심리학회지: 일반(2015.Vol 34)

학술논문

통합검색

단행본

전자책

학술논문

전자저널

연속간행물

학위논문

멀티미디어

소장기타



문화 창의

검색결과

☒ 전체

학술논문 10 / 10,977 건

숭실대학교 중앙도서관 학술논문 기준으로 두 키워드로 검색되는 논문 수는 약 11000건에 달할 정도로 많은 편이며

- ☐  중국 문화창의산업의 정책 패러다임 변동 연구 : 베이징시(北京市)를 중심으로 / Study on Policy Paradigm Change Framework on Cultural and Creative Industries in China : Focus on Beijing 阄
이건우 / Li Jian-yu
관련자료 KERIS dCollection (원문보기)
- ☐  코로나19가 중국 문화창의산업에 미치는 영향과 발전방안에 관한 연구 / A Study on the Influence and Development of COVID-19 on China Culture and Creativity Industry 阄
손소함 / SUN XIAOHAN
관련자료 KERIS dCollection (원문보기)

중국의 문화창의 정책을 시작으로 아시아권 국가들에서 창의문화 or 문화창의 라는 이름의 정책이 다수 시행되고 있다(대한민국의 경우에는 제주도에 창의문화캠퍼스 운영중).

<창의성과 취업*인재 관계> - 김성윤

- 창의성 양성과 관련되어 진행되고 있는 수많은 프로그램들이 창의성의 중요성을 대변



"창의" 프로그램



Q 전체

이미지

뉴스

동영상

쇼핑

더보기

도구

검색결과 약 20,400개 (0.24초)

신아일보

2023년 창의융합프로그램 '우리 같이 시보드게임 할래?' 참가자 ...

'참가자를 모집한다. 홍천군청소년수련관은 이달 3일부터 15일까지 관내 초등학교 4~6학년 20명을 창의융합프로그램 '...

1일 전



세계환경신문

'어린이가 행복한 전라북도' 대축제 성황리 개최

전라북도는 제101회 어린이날을 맞아 5일 전라북도 어린이창의체험관에서 어린이날 ...
낭독을 비롯해 온가족이 즐길 수 있는 다양한 체험놀이 프로그램으로 진행됐다.

10시간 전



"창의"



Q 전체

이미지

동영상

뉴스

지도

더보기

도구

검색결과 약 39,000개 (0.24초)

세계타임즈

군서미래국제학교 10학년 창의제안팀프로젝트 시흥시의의회와 ...

◦ '창의제안팀프로젝트'의 결과물을 기반으로 시흥시 창의체험학교 시흥청소년모의의회 프로그램과 연계하여 5월 2일에 제안서를 작성하고 모의의회에서 시의원이 되어...

15시간 전



내외통신

창의행정의 닳을 올리다...도봉구, MZ 세대 중심 아이디어 ...

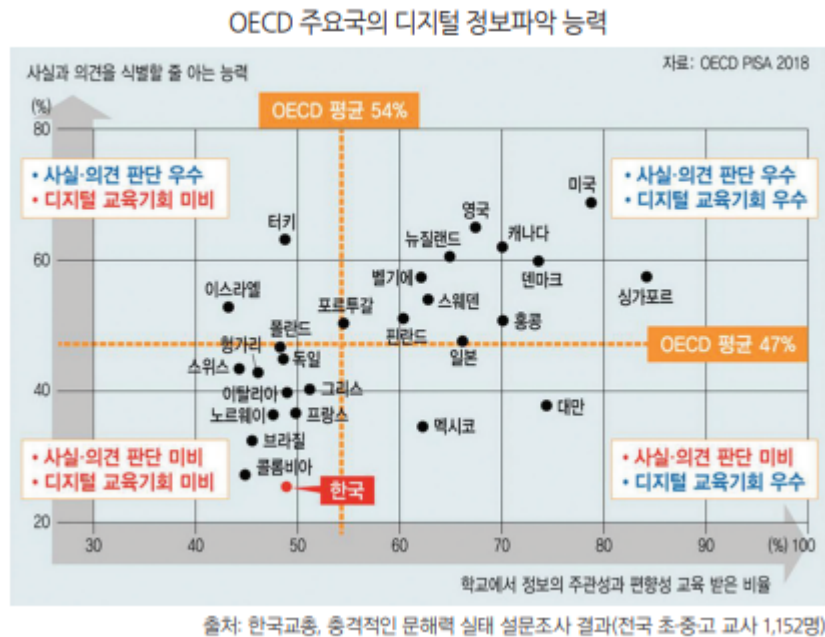
더불어 구는 올 하반기에 참여직원의 아이디어 발굴 역량을 강화하기 위해 제안 전문 강사를 초빙하여 동아리 회원들에게 창의행정, 창의제안 교육을 진행할 예정이다.

3일 전



(지난 1년간의 '창의' 혹은 '창의 프로그램'이 포함된 뉴스 검색결과 수가 상당히 많은 것을 알 수 있다. 높은 단어 빈도 수를 기반으로 사회가 생각하는 창의성의 중요성을 알 수 있다.)

- OECD에서 진행한 창의성 조사를 통해서도 알 수 있음



(한국의 문해력 문해력은 굉장히 낮은 수준, 출처 :

<https://www.hangyo.com/news/article.html?no=98072>)

“많은 전문가들은 ‘독서 부족’을 지적합니다. 문해력은 독서를 통해서 높일 수 있는데, 디지털 시대와 스마트폰의 일상화가 독서를 멀리하도록 만들었다는 것입니다.”

(출처 : <https://www.newstof.com/news/articleView.html?idxno=11797>)

“지식 정보처리 역량이 밑바탕이 되어야 창의적 사고를 키우고, 의사소통 역량을 기를 수 있는 것이다.” (출처 :

<https://news.edupang.com/news/article.html?no=12655>)

“한편, AI를 통해 생산성은 높이고, 인간 작업자는 창의성, 관리, 비판적 사고를 포함한 더 중요한 작업에 집중하는 등 노동의 총 수요에 영향을 미치지 않는 방향으로 적용 가능” (출처 :

https://hrstpolicy.re.kr/kistep/kr/board/BoardDetail.html?board_seq=53484&board_class=BOARD03&rootId=2006000&menuId=2006101)

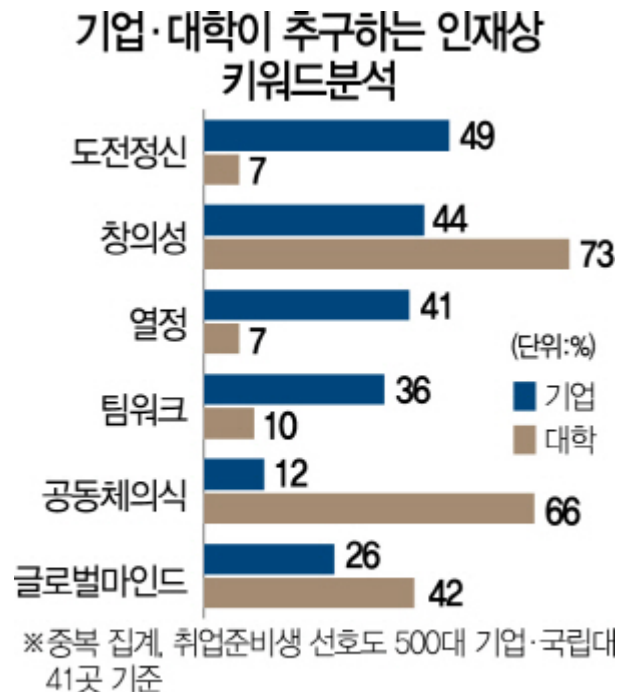
- 많은 기업들의 평가 요소로 창의성이 이용되는 점에서 기업에서도 창의성을 중요하게 여김을 알 수 있음

P : 대학은 창의성, 융합, 인성을 높은 순위의 인재상으로 뽑고 있지만, 실제로 기업에서는 창의성에 대해서 큰 중요도를 두고 있지 않은 모습을 보여줌

(<https://news.unn.net/news/articleView.html?idxno=210846>)

(<https://www.c1news.kr/news/articleView.html?idxno=40030>)

S : 이런 문제점을 지적하신다면, 창의성은 꾸준히 상위권에 존재했었다. 최상위 순위의 경우 시대상이 반영되는 경우이고 일시적인 영역일 수 있다.
(시대상 반영이라는 근거 자료,
<http://www.edujin.co.kr/news/articleView.html?idxno=38200>)



자료: 각 기업 홈페이지·대학교육협의회·진학사 캐치
(이 자료의 본 목적은 대학과 기업의 인재상의 미스매치지만, 대학과 기업이 모두 창의성을 높게 평가한다는 점을 제시하는데 활용할 수 있어보임, 출처 :
<https://m.sedaily.com/NewsView/1S8M4VWRBC>)

<표2> 업종별 인재상 순위

구분	제조업	금융·보험 업	무역·운수 업	건설업	도소매업	기타서비스 업
1위	도전정신	원칙신뢰	책임의식	소통협력	책임의식	책임의식
2위	책임의식	도전정신	도전정신	도전정신	소통협력	소통협력
3위	소통협력	책임의식	소통협력	원칙신뢰	전문성	창의성
4위	창의성	소통협력	창의성	책임의식	열정	열정
5위	열정	창의성	전문성	창의성	도전정신	사회공헌
6위	원칙신뢰	전문성	열정	전문성	원칙신뢰	원칙신뢰
7위	전문성	열정	글로벌역량	글로벌역량	실행력	글로벌역량
8위	실행력	글로벌역량	원칙신뢰	열정	창의성	전문성
9위	글로벌역량	사회공헌	실행력	실행력	사회공헌	도전정신
10위	사회공헌	실행력	사회공헌	사회공헌	글로벌역량	실행력

(100대 기업의 인재상 순위 변화표, 마찬가지로 창의력이 낮아짐을 대표하기 위한 자료지만, 아직 10위안에 든다는 점을 강조, 출처 :

<https://m.ekn.kr/view.php?key=20230130010006302>)

- 문학

P : 오히려 문학 도서 판매는 증가함....

<에스24 2018~2020년 연도 별 어린이, 청소년 문학 도서 판매 권수 및 판매 증감률>

기간 (2/23~3/15)	어린이 문학		청소년 문학	
	판매 권수 (일의 자리 반올림)	전년대비 증감률	판매 권수 (일의 자리 반올림)	전년대비 증감률
2018년	72,710	-	22,070	
2019년	96,570	32.8%	25,800	16.9%
2020년	158,340	64%	50,680	96.4%

*어린이 도서의 경우, 어린이 문학 내 동화책 카테고리 판매량을 기준으로 함

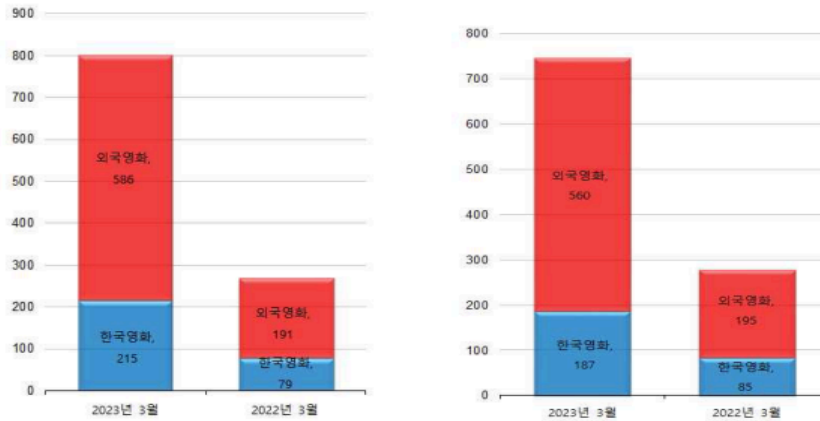
<분석데이터> - 이성일

<표 3> 2019년, 2021~2023년 3월 개봉편수, 매출액, 관객 수, 점유율 비교

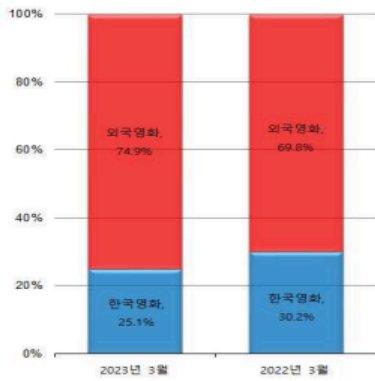
월	구분	2023년				2022년				2021년				2019년			
		개봉 편수	매출액 (억 원)	관객 수 (만 명)	관객 점유율 (%)	개봉 편수	매출액 (억 원)	관객 수 (만 명)	관객 점유율 (%)	개봉 편수	매출액 (억 원)	관객 수 (만 명)	관객 점유율 (%)	개봉 편수	매출액 (억 원)	관객 수 (만 명)	관객 점유율 (%)
1	한국	43	449	446	39.7	55	220	227	39.8	65	12	14	7.8	36	996	1,195	65.9
	외국	81	791	679	60.3	83	336	344	60.2	65	146	165	92.2	99	515	617	34.1
	소계	124	1,241	1,125	100.0	138	556	572	100.0	130	158	179	100.0	135	1,512	1,812	100.0
2	한국	32	134	127	19.8	43	127	138	42.1	68	61	68	21.9	22	1,463	1,723	77.3
	외국	72	556	515	80.2	69	182	189	57.9	81	226	243	78.1	81	437	505	22.7
	소계	104	691	642	100.0	112	309	327	100.0	149	287	311	100.0	103	1,900	2,228	100.0
3	한국	52	215	187	25.1	54	79	85	30.2	92	35	39	11.9	28	534	627	42.7
	외국	93	586	560	74.9	100	191	195	69.8	111	266	287	88.1	96	731	840	57.3
	소계	145	800	748	100.0	154	270	280	100.0	203	302	326	100.0	124	1,266	1,467	100.0
합계	한국	127	798	761	30.3	152	427	450	38.2	225	108	121	14.8	86	2,994	3,545	64.4
	외국	246	1,933	1,754	69.7	252	709	729	61.8	257	639	695	85.2	276	1,683	1,962	35.6
	합계	373	2,731	2,515	100.0	404	1,136	1,179	100.0	482	747	815	100.0	362	4,677	5,507	100.0

1. 코로나 유행 이전과 유행 중 , 유행 후 영화관 매출 및 관객 수 데이터

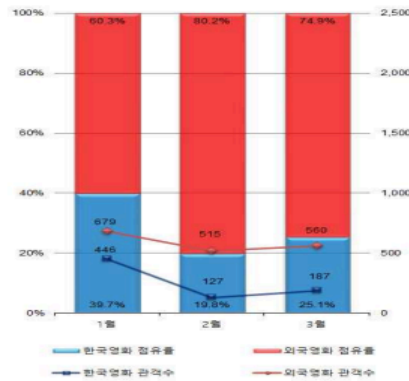
〈그림 1〉 2022~2023년 1~3월 매출액, 관객 수 비교
매출액 (단위: 억 원) 관객 수 (단위: 만 명)



〈그림 2〉 2022~2023년 3월 관객 점유율



〈그림 3〉 2023년 1~3월 관객 점유율 및 관객 수 (단위: 만 명, %)



2. 마스크 해제 등 포스트 코로나를 바라보는 2023년 영화 산업 관련 매출 지표

((영화진흥위원회) 출처:

<https://www.kofic.or.kr/kofic/business/board/selectBoardDetail.do?boardNumber=2#none>)

AREA_NM	DAIL_BOOK	WEEK_BOOK	MT_BOOK	QUARTER_BOOK
서울	4	14.4	20.8	28.9
부산	1	8.4	16.8	25
대구	3.7	12.3	23.3	16.5
인천	3.5	14.1	21.6	19.9
광주	2	10.7	11.8	21.1
대전	1.7	11.3	11.9	14
울산	0.4	11.6	14.3	27.1
세종	4.8	10.6	13.1	19.7
경기	1.9	11.7	14.8	26.5
강원	5.8	13.2	18.5	17.5
충북	1.5	12.1	14.8	16.3
충남	0.6	5.6	15.1	17.9
전북	0.7	10.6	6.3	17.6
전남	1.9	7.6	16.3	14.3
경북	2	9.9	6.8	10.5
경남	0	4.1	14.7	20.6
제주	3.6	19.4	11.7	19.9

3. 코로나 이전 독서율 (지역별)

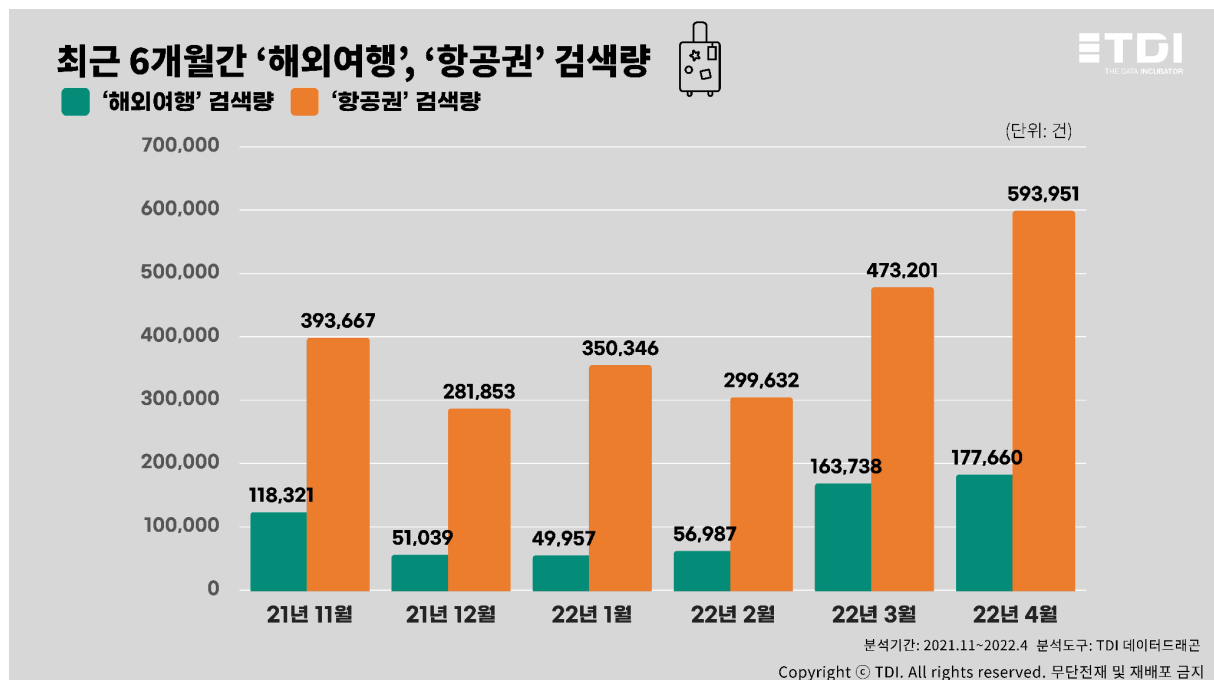
4. 고민이 필요한 코로나 시대의 도서관 문화

(출처: <https://www.iusm.co.kr/news/articleView.html?idxno=924256>)

5. 도서관 달히자 늘어난 전자책 대출

(출처:

[https://librarian.nl.go.kr/LI/contents/L30302000000.do?schFld=0&schStr=%EB%8F%84%EC%84%9C%EA%B4%80+%EB%8B%AB%ED%9E%88%EC%9E%90&schM=view&page=1&ordFld=regdt&ordBy=DESC&viewCount=9&id=5947&schBdcode=&schGroupCode=\)](https://librarian.nl.go.kr/LI/contents/L30302000000.do?schFld=0&schStr=%EB%8F%84%EC%84%9C%EA%B4%80+%EB%8B%AB%ED%9E%88%EC%9E%90&schM=view&page=1&ordFld=regdt&ordBy=DESC&viewCount=9&id=5947&schBdcode=&schGroupCode=)))



6. 여행 산업 회복되나...빅데이터로 살펴본 여행 업계

(출처: <https://mirakle.mk.co.kr/view.php?year=2022&no=406364>)

EXAMIN_BSEXDSTN	AGRDE_FL	ANSWRR_(HSHLD_IN	ALL_TURSI	DMSTC_TL	OVSEA_TURSI		
20230403	F	50대	부산광역시	300이상50	약간 더 많	약간 더 많	훨씬 더 적게
20230403	F	50대	서울특별시	700만원 0	훨씬 더 많	훨씬 더 많	훨씬 더 많이
20230403	F	50대	서울특별시	500이상70	약간 더 많	훨씬 더 많	훨씬 더 많이
20230403	F	50대	서울특별시	700만원 0	훨씬 더 많	약간 더 많	훨씬 더 많이
20230403	F	50대	서울특별시	300이상50	약간 더 많	약간 더 많	약간 더 적게
20230403	F	50대	서울특별시	500이상70	비슷하게	훨씬 더 적	비슷하게
20230403	F	50대	서울특별시	500이상70	비슷하게	훨씬 더 많	훨씬 더 많이
20230403	F	50대	서울특별시	500이상70	비슷하게	훨씬 더 적	약간 더 적게
20230403	F	50대	서울특별시	700만원 0	비슷하게	약간 더 많	약간 더 많이

7. 여행 비용 지출 의향 (2023) - cols = 시기, 성별, 나이, 거주지, 수입, 여행지출비용계획, 국내여행 -, 해외여행 -

■ 세계 관광산업 피해 실태

세계관광기구(UNWTO) 자료에 따르면, 코로나19의 영향으로 세계 관광시장은 30년 전 수준으로 회귀하였다. 세계 관광객 수는 약 10억 명이 줄어들면서 전년 대비 70~75% 감소세를 기록하였고 이로 인한 관광 수입의 감소 규모는 약 1조 1천억 달러에 달할 것으로 추산하였다. 코로나19로 위협에 처한 일자리는 약 10~12억 개이며, 이러한 영향을 종합했을 때 세계 GDP의 손실 규모는 2조 달러를 상회할 것으로 추산하였다.

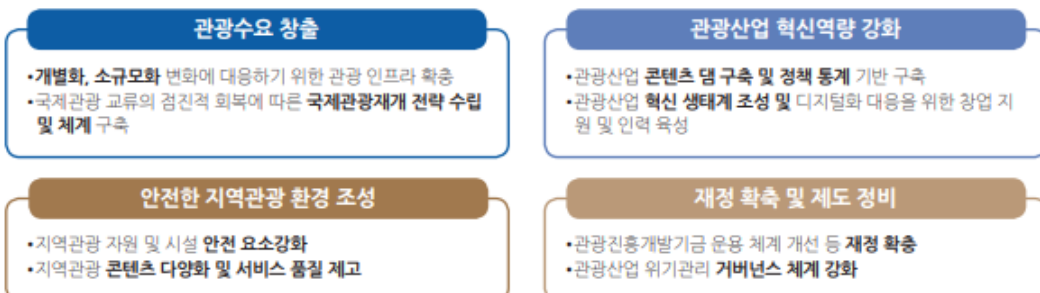
[그림 4] 코로나19로 인한 세계관광산업 피해 실태



자료: UNWTO(2020)

코로나19의 영향으로
세계 관광시장은 30년 전
수준으로 회귀

[그림 14] 코로나19 위기 극복을 위한 주요 과제



9. 관광산업 피해 실태 및 극복 방안 (출처:

https://policy.nl.go.kr/search/searchDetail.do?rec_key=SH2_PL20210261069)

10. 일본 관광산업의 부활, 새로운 관광산업 트렌드는 (포스트 코로나 시대 관광, 여행 산업 트렌드) (출처:

https://dream.kotra.or.kr/kotranews/cms/news/actionKotraBoardDetail.do?SITE_NO=3&MENU_ID=180&CONTENTS_NO=1&bbsGbn=243&bbsSn=243&pNttSn=200347)

EXAMIN_YSEXDSTN	AGRDE_FL	ANSWRR_H	SHLD_IN	LSTYR_VEF	LSTYR_VEF	TWO_YEA	TWO_YEAR_WITHIN
202203 M	30대	서울특별시 300이상50	"박물관, 영	"없음"	연극-뮤지	미술관-전시관	
202203 F	30대	서울특별시 500이상70	"미술관-전	"연극-뮤지	음악 공연	미술관-전시관	
202203 F	20대	서울특별시 무응답	"영화관, 음	"연극-뮤지	연극-뮤지	없음	
202203 M	60대	서울특별시 500이상70	"없음"	"없음"	영화관	없음	
202203 F	60대	서울특별시 500이상70	"음악 공연	"영화관"	음악 공연	없음	
202203 F	20대	서울특별시 300만원 0	"없음"	"없음"	연극-뮤지	음악 공연	
202203 M	60대	서울특별시 300만원 0	"없음"	"영화관"	연극-뮤지	없음	
202203 M	40대	서울특별시 700만원 0	"없음"	"박물관, 영	영화관	박물관	
202203 F	50대	서울특별시 500이상70	"영화관, 음	"없음"	문학 행사	영화관	
EXAMIN_YSEXDSTN	AGRDE_FL	ANSWRR_H	SHLD_IN	LSTYR_VEF	LSTYR_VEF	TWO_YEA	TWO_YEAR_WITHIN
202303 F	20대	서울특별시 300만원 0	박물관, 영	미술관-전	음악 공연	없음	
202303 M	20대	서울특별시 300만원 0	없음	없음	음악 공연	문학 행사	
202303 M	40대	서울특별시 700만원 0	없음	없음	음악 공연	연극-뮤지컬	
202303 M	40대	서울특별시 700만원 0	없음	없음	음악 공연	미술관-전시관	
202303 F	40대	서울특별시 500이상70	없음	연극-뮤지	연극-뮤지	미술관-전시관	
202303 M	40대	서울특별시 500이상70	없음	없음	연극-뮤지	없음	
202303 F	20대	서울특별시 무응답	영화관	없음	음악 공연	연극-뮤지컬	
202303 M	20대	서울특별시 무응답	없음	없음	음악 공연	무용 공연	
202303 F	60대	서울특별시 300이상50	미술관-전	연극-뮤지	문학 행사	미술관-전시관	

11. 2022와 2023 문화 예술 관람 현황 및 희망 분야 - cols= 기간, 성별, 나이, 거주지, 수입, 작년 대비 증가 문화활동, 작년 대비 감소 문화 활동, 1순위 희망 문화 활동, 2순위 희망 문화활동

2019	1	서울특별시영유아(0-5)남자	20962	2034	10.306
2019	1	서울특별시영유아(0-5)여자	18840	1810	10.409
2019	1	서울특별시유아(6-7) 남자	36922	3533	10.451
2019	1	서울특별시유아(6-7) 여자	32749	3061	10.699
2019	1	서울특별시초등(8-13) 남자	155868	17424	8.946
2019	1	서울특별시초등(8-13) 여자	152671	16834	9.069
2019	1	서울특별시청소년(14-남자	43580	9194	4.74
2019	1	서울특별시청소년(14-여자	53337	11646	4.58
2019	1	서울특별시20대 남자	39760	10544	3.771
2019	1	서울특별시20대 여자	71240	18424	3.867
2019	1	서울특별시30대 남자	102891	17260	5.961
2019	1	서울특별시30대 여자	242090	37049	6.534
2019	1	서울특별시40대 남자	263440	36725	7.173
2019	1	서울특별시40대 여자	388380	56159	6.916
2019	1	서울특별시50대 남자	83387	14206	5.87
2019	1	서울특별시50대 여자	77567	15067	5.148
2019	1	서울특별시60대이상 남자	64382	10265	6.272
2019	1	서울특별시60대이상 여자	44742	7039	6.356

2020	1	서울특별시영유아(0-5)남자	17653	1859	9.496
2020	1	서울특별시영유아(0-5)여자	19140	1851	10.34
2020	1	서울특별시유아(6-7) 남자	32144	3159	10.175
2020	1	서울특별시유아(6-7) 여자	30209	2988	10.11
2020	1	서울특별시초등(8-13) 남자	145716	16339	8.918
2020	1	서울특별시초등(8-13) 여자	157863	17518	9.011
2020	1	서울특별시청소년(14-남자	36484	7555	4.829
2020	1	서울특별시청소년(14-여자	50821	10947	4.642
2020	1	서울특별시20대 남자	29656	7981	3.716
2020	1	서울특별시20대 여자	66379	17870	3.715
2020	1	서울특별시30대 남자	73790	12853	5.741
2020	1	서울특별시30대 여자	232433	36362	6.392
2020	1	서울특별시40대 남자	234038	32123	7.286
2020	1	서울특별시40대 여자	421370	61944	6.802
2020	1	서울특별시50대 남자	78528	13968	5.622
2020	1	서울특별시50대 여자	84782	16972	4.995
2020	1	서울특별시60대이상 남자	62827	10340	6.076
2020	1	서울특별시60대이상 여자	48432	7822	6.192

12. 코로나 유행 전과 유행 중일 때 독서량 변화 – cols = 년도, 분석달, 지역, 연령, 성별, 대출 수, 대출 멤버십 수 , 독서량

<데이터 리스트>

- [연령별 도서 키워드 데이터](#)
- [연령별 특정도서 데이터](#)
- 권장도서 대출현황 분석
- 학급별 사교육비 대비 대출분석
- 도서별 상세정보
- 공공 도서관 대출정보
- [공공도서관 소장 데이터](#)
- 공공 도서관 소장도서

- 도서 감성 데이터
- 인기 대출
- 연간 독서량(율)
- 대출예약 도서순위
- 도서관별 대출추이 분석
- 도서 소장 순위
- 지역별 장서회전율
- 세트 **ISBN** 데이터 (**ISBN** : 도서 고유 넘버)
- 신착도서
- 베스트 셀러
- 대출상승 도서목록
- **yes24** 베스트 셀러 데이터
- 교보문고 베스트 셀러 데이터
- kosis의 전자책 데이터
- 공공도서관 데이터
- 서울시
독서실태조사(<http://data.seoul.go.kr/dataList/OA-15556/F/1/datasetView.do>)

<데이터로부터 아이디어> 14가지

1. **도서별 특정 키워드 유행주기가 있나?**
2. **웹소설의 인기 -> 도서관에서는 그런 책을 유치하지않음 왜?**
3. **웹소설 -> 웹툰 -> 드라마화된 것들의 공통점 특징점 파악 -> 해당 요소 극대화 (위키 데이터)**
4. **홍보만 잘하면 장땡? 의미없는 도서도 인기순위가 되는 마법 -> 해당 도서들의 공통점**
이슈: 해당 도서에 대한 부정적 의견을 어떻게 파악가능?
5. **인기없는 도서들을 도서관에서 가지고 있는 이유?**
6. **도서관의 이용 도서, 서점의 구매도서 공통점 차이점 분석 -> 도서관 이용률 증대 방안 제시가능**
7. **키워드 분석을 통해 장르별 인기있는 키워드 추출 가능**
8. **도서별 높은 순위에 위치한 이유를 분석(제목, 작가, 장르, 높은 순위에 머무른 정도)**
9. **도서관 대출이 증가한 도서들의 특징을 분석(도서 정보를 기반)**
- 10.
11. **도서관별 소장도서의 공통점을 분석 -> 도서관들의 도서 선정 기준 유추**
12. **도서관별 대출 추이를 통해 대출이 많이 발생하는 시기 예측 혹은 분석**
- 13.
14. **영상 매체 (유튜브 등)을 많이 이용하고 책을 적게 읽을 시 실제로 문제가 있는지 분석을 통해 인식 제고**

<아이디어 조합> 6가지

- **2, 5, 11번은 합치기**
- **9, 12번도 합치기**
- **14번 + 교육 데이터 활용 good**
- **4, 7, 8번 합치기**
- **1,6 번 합치기기**
- **3번**

기대효과?에 대해 설명

<아이디어 조합의 기대효과>

1. 도서관뿐 아니라 서점, 다양한 매체에서도 베스트 셀러가 될 수 있는 방안 제시 가능, 인기도서들의 제목 분석가능
2. 대출 분석을 통해 도서관 측이 소장할 도움을 줄 수 있다.
3. 이슈 : 실제로 문제가 있다 -> 엠지세대의 문해력 문제는 진짜인가?
4. 어떤 책이 인기가 있는 것인지 키워드, 제목, 장르 등 다각적으로 분석
5. 서점에서의 유행을 분석하여, 도서관의 유행주기도 예측하여 도서관 측에는 도서를 소장할 때, 도움이 될 수 있다.
6. 우리나라의 문해력 강화가능 매출증대 -> 다시 새로운 콘텐츠 개발 -> 선순환 포장했지만 드라마화되고 웹툰화되어서 돈 많이 벌면 당연히 좋은거 아닙니까
이슈 : 데이터 부족.. 제목만 알 수 있음.

<아이디어 찌막 조합>

1. 1, 2, 4, 5번 합치기(1, 2, 4, 5, 6, 7, 8, 9, 11, 12 번) -> 도서관의 새로운 도서 선정의 도움 받기 -> 도서관 이용률 증대 -> 창의 인재 형성 -> 국가 발전 -> 선진강국 대한민국 !!!!! 사랑해요 연애가 중계~

교수님 예상 태클?

2. 3번
3. 6번
- 4.

<활용할 아이디어>

주제 :

- 도서별 특정 키워드 유행주기가 있나?
- 웹소설의 인기 -> 도서관에서는 그런 책을 유치하지않음 왜?
- 홍보만 잘하면 장땡? 의미없는 도서도 인기순위가 되는 마법 -> 해당 도서들의 공통점
이슈: 해당 도서에 대한 부정적 의견을 어떻게 파악가능?
- 인기없는 도서들을 도서관에서 가지고 있는 이유?
- 도서관의 이용 도서, 서점의 구매도서 공통점 차이점 분석 -> 도서관 이용률 증대 방안 제시가능
- 키워드 분석을 통해 장르별 인기있는 키워드 추출 가능
- 도서별 높은 순위에 위치한 이유를 분석(제목, 작가, 장르, 높은 순위에 머무른 정도)
- 도서관 대출이 증가한 도서들의 특징을 분석(도서 정보를 기반
- 도서관별 소장도서의 공통점을 분석 -> 도서관들의 도서 선정 기준 유추
- 도서관별 대출 추이를 통해 대출이 많이 발생하는 시기 예측 혹은 분석

-> 도서관과 서점에서 인기있는 책 장르, 제목 선정하기

궁극적으로 목표하는 바가 무엇인가????

-> 문해력 감소 연구한 논문많음(배경, 목표) -> 도서관 사용률 증대측면(분석 데이터) -> 해당 도서들을 비치시켜도 도서관 사용률이 증가되는가를 입증가능??

인기있는 책이 있다고 사용률이 증대될 수 있는가? -> 기존 사용자들의 단순히 만족도를 높이는 방향으로

-> 기존의 책들의 데이터 분석을 통한 “도서관 소장 도서 추천(= 인기 있을 도서 추천)” 이건 별론 가용

“찾은 데이터 조합을 했더니 너무 흔한 주제이다”

-> 책의 매출 증대! 국가적 구매 -> 개꿀

-> 좋은 책, 추천할만한 책 선정 가능

-> 어그로를 원활히 끌만한 제목 선정 가능

<주의할 점>

- 기업과 달리 고객 자체의 데이터를 보유하지 않으므로 배제해야함.
- 한 데이터셋을 기반한 단순한 통계 & 모델만 쓰면 나오는 결과는 배제해야함.
- 목표와 데이터 분석 결과를 매칭해야함.
- ***자신감의 잃지 않아야함***

<5/15일자 업데이트>

배경(00을 보니 00의 의문을 생각하게 되었다)

- 엠지세대의 문해력 부족 문제 대두
- 코로나19로 인해 도서관이 도서 박물관으로서 보이는 양상 관측 -> 도서관 도서 사용량 증대 방안 모색

목표(이 의문을 해결하고자 00을 분석해보고자 한다)

- 도서관 데이터를 활용한 도서 추이 파악과 원인 분석

기대효과(이 분석을 한다면 000한 기여를 할 수 있다)

- 다수가 만족하는 도서를 준비해서 도서관 이용률을 높인다.
- 도서관 예산의 효율적 사용

<도서관 소장 데이터 && 도서관 대출정보 => 도서관 타겟 분석 & 이유 파악>

1. 코로나 전 후 비교 (데이터 간격은 진행하면서 정하기 -> 분석 기법 중 하나)
2. 소장 vs 대출 => if yes (둘 중 하나의 데이터로 이유 분석) else (그 간격에 대해서 파악) + 키워드 데이터
3. 실제 판매가 많이 되는 자기계발서, 주식투자 많은가? -> 도서관은 이 사실을 알고 반영을 하는가?
4. (가능하면) 서점 데이터하고도 비교 -> 서점에서 베스트셀러로 판매되는 책들이 도서관에 충분히 들어오는가? yes? no?

발생 가능한 이슈(모든 이슈를 해결할 수 없다. 단지 생각을 충분히 해보았는가에 초점을 맞춘다):

1. 도서관의 주 목표가 다양한 지식 파악인데 트렌드 도서들을 채우는게 맞는 것인가

현재까지 진행상황 정리

1. 데이터 확정 (도서관 소장 도서 데이터, 도서관 대출정보)

교수님께 질문 :

1. 배경을 가지게 된 계기에 대해 증명하지 않아도 되는가?

결정해야하는 내용 :

1. 이 데이터를 예측 vs 분석 => 분석 쪽으로 진행할 예정
2. 사용할 데이터 => 대출정보, 소장도서(=ㅇ 많음)
3. 사용할 데이터 주기 => 1년 주기
4. 데이터별 컬럼 정하기 =>

<데이터 리스트>

- [연령별 도서 키워드 데이터](#)
- [연령별 특정도서 데이터](#)
- 권장도서 대출현황 분석
- 학급별 사교육비 대비 대출분석
- 도서별 상세정보
- [공공 도서관 대출정보](#)
- [공공도서관 소장도서](#)
- 도서 감성 데이터
- 인기 대출
- 연간 독서량(율)
- 대출예약 도서순위
- 도서관별 대출추이 분석
- 도서 소장 순위
- 지역별 장서회전율
- 세트 **ISBN** 데이터 (**ISBN** : 도서 고유 넘버)
- 신착도서
- [베스트 셀러](#)
- 대출상승 도서목록
- **yes24** 베스트 셀러 데이터
- 교보문고 베스트 셀러 데이터
- [kosis의 전자책 데이터](#)
- 공공도서관 데이터
- [각 도서관별 대출 데이터](#)
- [인기대출](#)

하나의 핵심질문?

장르? 연령대별 성별에 따른 차이, 관련정보에 따라 달라진다
제안발표때 중요하게 생각한.. 왜 중요하고 왜 해야하는가?

질문이 약한가? -> 배경 목표가 약하다
데이터를 늘리는 방안

중간발표는 초기결과를 제시하여야 한다.

목요일

<도서관 소장도서>

순 서	컬럼영문명	컬럼한글명	데이터타 입	길이	PK여 부	NOT NULL여부
1	CTRL_NO	제어번호	varchar	20	Y	Y
2	AUTHR_NM	저작자명	varchar	4000	N	N
3	VLM_NM	권명	varchar	150	N	N
4	PBLICTE_YEAR	발행년도	varchar	100	N	N
5	CL_SMBL_NO	분류기호번호	varchar	22	N	N
6	BOOK_SMBL_NO	서적기호번호	varchar	22	N	N
7	TITLE_NM	제목명	varchar	1000	N	N
8	LBRRY_CD	도서관코드	int	1	Y	Y
9	ISBN_THIRTEEN_NO	ISBN13번호	varchar	13	N	N
10	REPRSNT_BOOK_AT	대표서적여부	char	1	N	N
11	REGIST_NO	등록번호	varchar	12	N	N
12	INCME_FLAG_NM	수입구분명	varchar	10	N	N
13	MANAGE_FLAG_NM	관리구분명	varchar	10	N	N
14	MEDIA_FLAG_NM	미디어구분명	varchar	10	N	N
15	UTILIIZA_LMTT_FLAG_NM	이용제한구분명	varchar	10	N	N
16	UTILIIZA_TRGET_FLAG_NM	이용대상구분명	varchar	10	N	N
17	ACP_DATA_NM	동반자료명	varchar	200	N	N
18	SGVL_ISBN_NO	날권ISBN번호	varchar	50	N	N
19	SGVL_ISBN_ADTION_SMBL_N M	날권ISBN부가기호 명	varchar	60	N	N
20	CL_SMBL_FLAG_NM	분류기호구분명	char	1	N	N
21	VLM_SMBL_NM	권기호명	varchar	100	N	N
22	DUCACPY_SMBL_NM	복사본기호명	varchar	22	N	N
23	REGIST_DE	등록일자	datetime		N	N
24	ISBN_THIRTEEN_ORGT_NO	ISBN13원문번호	varchar	13	N	N
25	MASTR_LBRRY_CD	마스터도서관코드	int		N	N
26	VLM_EXST_AT	권존재여부	char	1	N	N
27	SET_ISBN_CHG_AT	집합ISBN변경여부	char	1	N	N
28	VLM_ORGT_NM	권원문명	varchar	150	N	N

29	TITLE_SBST_NM	제목대체명	varchar	200	N	N
30	KDC_NM	KDC명	varchar	22	N	N
31	BOOK_CL_CD	서적분류코드	varchar	20	N	N
32	BOOK_LC_CD	서적위치코드	varchar	20	N	N

<도서관 대출 정보>

순서	컬럼영문명	컬럼한글명	데이터타입	길이	PK여부	NOT NULL여부
1	SEQ_NO	일련번호	varchar	100	Y	Y
2	BOOK_KEY_NO	책제어번호번호	varchar	20	N	N
3	MBER_SEQ_NO_VALUE	회원일련번호값	varchar	100	N	N
4	LON_DE	대출일자	datetime		N	N
5	RTURN_DE	반납일자	datetime		N	N
6	LBRRY_CD	도서관코드	int	1	Y	Y
7	LON_STLE_NM	대출형태명	varchar	10	N	N
8	RTURN_STLE_NM	반납형태명	varchar	10	N	N
9	RTURN_PREARNGE_DE	반납예정일자	datetime		N	N
10	RESVE_DE	예약일자	datetime		N	N
11	RESVE_END_DE	예약만료일자	datetime		N	N
12	LON_STATE_NM	대출상태명	varchar	10	N	N
13	LON_PLACE_NM	대출장소명	varchar	50	N	N
14	RTURN_PLACE_NM	반납장소명	varchar	50	N	N
15	MANAGE_LBRRY_CD	관리도서관코드	varchar	50	N	N
16	LON_LBRRY_CD	대출도서관코드	varchar	50	N	N
17	RTURN_LBRRY_CD	반납도서관코드	varchar	50	N	N
18	MASTR_LBRRY_CD	마스터도서관코드	int		N	N

필요한 데이터가 뭐였죠

제목별로 몇 권을 가지고 있었는가?

<도서관 소장 데이터 && 도서관 대출정보 => 도서관 타켓 분석 & 이유 파악>

코로나 전 후 비교 (데이터 간격은 진행하면서 정하기 -> 분석 기법 중 하나)

소장 vs 대출 => if yes (둘 중 하나의 데이터로 이유 분석) else (그 간격에 대해서 파악) + 키워드 데이터

실제 판매가 많이 되는 자기계발서, 주식투자 많은가? -> 도서관은 이 사실을 알고 반영을 하는가?

(가능하면) 서점 데이터하고도 비교 -> 서점에서 베스트셀러로 판매되는 책들이 도서관에 충분히 들어오는가? **yes? no?**

목표? 기대효과를 확실히 하고 넘어가야..

할 일

1. 분류 코드 분석을 통한 장르 구분 -> 도서관측 요청?
2. 대출정보 누적을 활용해서 대출 인기도 유추? (베리곤)
3. 시기별 소장도서를 (저희만의 기준)으로 정리 => 시기별 변화를 파악하자
4. 서점 베스트 셀러를 바탕으로 해당 시기에 해당도서가 도서관에 있는지 파악
- 5.

영민 - 목표와 기대효과를 정하고 데이터를 맞춰 나아가자 (현재 살짝 불분명하다)

호건 - 목표와 기대효과를 정하자 (데이터들은 어느정도 이제 확보가 되었다)

성일 - 특정 도서관을 정해서 소장도서와 대출을 비교한다.(그래도 목표와 기대효과를 확실히 정해야한다)

성운 - 목표와 기대효과를 정하긴 해야하지만 틀에 갇힐 필요는 없다 (확실한 방향성만 잡힌다면 세부적인 것은 상황에 맞게 적용)

목표 :

기대효과 :

나이, 성별, 기간 중 택1

지난 기대 효과 : 분석을 통해서 변화의 추이를 보고 도서관 예산을 효율적으로 운용하는 것 효율적으로 운영하려면?

성일 : 나이 -> 장르 대출 추이 분석 -> 도서관 입장에서 특정연령 타게팅에 쓰일 전략 제시

호건 : 성별 -> 장르 대출 추이 소장 비율과 비교 분석 -> 왜 특정 성별의 이용 비율이 높을까?

영민 : 기간 -> 쓸데없는 인기없는 책에 많은 돈을 쏟는가? 일정 기간별로 새로 들어온 책에 대한 대출량을 보고 거기에 예산을 사용하고 있나..

성운 : 기간 -> 기간별 인기 도서의 변화를 살피고 그 추이를 해석해서 사회 문제(해석을 통계자료... 데이터를 찾아서 비교해보고 상관관계 파악해보면서 되지 않을까 싶습니다.

특정 시간에 대한 상관관계)를 시사하는 방향으로

대외적인 이슈와 관련 어떤거든 (변화를 보고 그 이유를 사회문제에서 찾는 느낌으로)

예, 조국 사건을 터졌을 때, 조국 자서전.. 지지자들의 의해 매출이 괜찮았다. 이걸

사회적으로 봤을 때는 문제를 일으킨 사람인데, 인기 도서가 됐다.)

기간별 특정 주제별 대출권수를 파악해서 이 부분에서 뭔가 **outlier**가 발견되면 그 지점을 좀더 파헤쳐 보는 식으로...

어떤 데이터? <도서관 정보나루>

기간별, 성별, 나이별, 지역별, 주제별 데이터가 있음

1. 데이터 선정 -> 10시

- 기간 데이터 / 장르 데이터 / 대출량 데이터
- => 도서관 정보나루 활용
- => 중간 발표 : 장르-경제
- => 기간 : **2019~현재 (19(성윤), 20(호건), 21(영민), 22(성일))**-> 최소 월별로 하나씩은 가져오기
- => 이슈 기준(뉴스*기사 보도) 앞뒤 주기 : **2달**
- => 이슈 선정 대응 : 가장 영향이 있어보였다. 가설 검증.
- => 각자 맡은 기간별 이슈 찾아오기

2. 데이터 전처리 해서 인기 도서 변화 -> 목요일 미팅 전까지 '방향'만은 꼭 생각해오기

3. 주제별 대출수 특징 파악(갑자기 대출이 늘어난 장르?) -> 목요일 미팅

4. 특징과 관련된 이슈, 원인을 찾아낸다 (= 연결짓는다) -> 목요일 미팅, 1개 정하기

5. 결과 모음 => 기대효과를 가져감 (+배경) -> 카톡 혹은 온라인 미팅

6. 발표 자료 발표 내용

회의록! -> 오늘은 호건님!

<숙제>

각자 맡은 기간 내에 이슈 찾아오기 (수요일 저녁 전까지)

이슈에 맞는 데이터 찾기 (최소 월별 1개) (목요일 미팅 전까지)

• 영민	→ 21
• 성윤	→ 19
• 호건	→ 20
• 성일	→ 22

중간 : 한가지 특징만이라도 납득할 수 있는 연결다리를 만들어냈으면 좋겠습니다.

예, 조국. 박근혜. 이명박 사건을 통해 ~경향이 있는 것을 알 수 있었다.

예, 경제쪽으로 선정 (누가 주식으로 돈을 많이 벌었다라는 이슈 -> 주식책 증가)

예, 큰 사고(이태원, 세월호) 우울감이 미치는 영향

기말 : 둘 이상의 특징으로

-> 기대효과: 특정 장르 도서의 출판일 결정 가능 & 도서관에 해당 장르 도서 더 비치하는 정책 제시

엠쥬놈들 문해력 딸림 -> 증대방안 중 하나는 책 많이 읽기 -> 대표적인 것이 도서관 사용량 증대

도서관을 도서관박문관으로서 사용하는 모습 독서실을 차라리 만들지 그럼

1. 도서관 사용량 증대 -> 인기도서 구비
2. 쓸데없는 인기없는 책에 많은 돈을 쏟는가?
3. 인기 있는 장르가 시기별로 존재한다 -> 장르 추천

배경 - 문해력 문제 등 -> 도서관 이용 활성화

목표효과 -> 기간별 대출 데이터를 비교해서 장르, 연령 분석

후 -> 인기있는 책이 무엇인지 확인 -> 도서관 이용 활성화를 도울 수 있는 방안 제시

2019년 주요 경제 (김성윤)

1월 :

2019년 첫 금통위 “기준금리 동결, 성장률전망 0.1%p 하향”

美, 경기 하강세 뚜렷...G2 무역전쟁 장기화 ‘최대복병’

[미국 경제마저 꺾이나?] 10년 ‘나홀로 호황’ 변곡점 다다를 듯

2019년 사회적경제예산 474억...박원순시장 “사회적기업 ‘규모화’ 돕겠다”

2월 :

거대한 ‘빛’에 포위된 2019년 세계경제

경기도, 2019년 경제정책 키워드는 ‘공정경제와 혁신성장’

2019년 국내 은행 전망... ‘가시밭길 예상’

3월 :

미국 경제의 이슈가 된 부유세 논란

‘ICT 아닌 게 없다’ 2019년에도 4차산업혁명 속도낸다

4월 :

2019년 양회를 통해 본 중국의 경제정책 방향 및 시사점

5월 :

일본 경제: 예상 깨고 ‘깜짝 성장’

6월 :

식지 않는 경제 성장을 논란

UNWTO “2019년 세계 여행객 15억명 돌파”

7월 :

한일 무역 분쟁

8월 :

블룸버그 “한국 주식시장, 2019년 최악”평가

9월 :

금융 불안정, 부채 위기: 세계경제 또 위기 직면

10월 :

개혁·개방 가속화하는 중남미

비정규직 87만명 급증...’숨은 50만명’ 드러났다

11월 :

한국 국가경쟁력 세계 13위

“고장난 자본주의를 고쳐라” 대안 모색하는 세계

12월 :

저성장·저물가에 신음하고 서울 집값은 폭등

"44% 오른 삼성전자, 아직 저평가"...PER, 글로벌경쟁사의 절반

한중평 : 미vs중, 한vs일 => 전체적으로 경제 상황이 좋지 못했다.

1. 한일 무역분쟁
2. 미중
3. covid 개막

2020년 주요 사건 (이호건)

- 1월 : 영국 브렉시트
- 2월 : 코로나 경보 심각으로 격상 / 영화 기생충 아카데미 시상식 4관왕
- 3월 : 코로나로 인한 경제 타격 대비 금리 인하(첫 0%대 금리)
- 4월 : 제21대 국회의원 선거
- 5월 : 조지 플로이드 사망 사건
- 6월 : 북한 남북공동연락사무소 폭파
- 7월 : 주택 임대차보호법 시행
- 8월 : 의료정책 추진 반대 시위
- 9월 : 방탄소년단 한국 최초 빌보드 1위
- 10월 : 도쿄증권거래소 오류로 인한 거래 중지
- 11월 : 미국 대통령 선거
- 12월 : 코로나19 백신 승인

2021년 주요 경제 (강영민)

1월 :
한국이 체결한 최초의 다자·메가 FTA
바이든 시대 & 미국 의회 습격사건
LG 전자 스마트폰 철수

2월 :
코로나 위기의 온전한 극복과 선도형 경제
지역균형 뉴딜 전략 추진 등 지역경제 회복과 균형발전
미 - 중 분쟁 심화

3월 :
탄소중립
코스피 3000 돌파
NFT와 디지털자산의 시대

4월 :
한국 스타트업의 중흥기
행복주택 및 주거 문제
백신여권의 등장
일반인 백신접종 시작 다가오는 집단면역
가파른 부동산 가격 상승 & 근로 의욕 상실

5월 :

[수소경제](#) 및 수소차
[한미 정상회담](#)

6월:

[미국 최악의 구인난](#)
[재택 근무 vs 출근](#)

7월:

[도쿄 올림픽 개막](#)
[한·EU FTA 10주년](#)

8월 :

[개인형 이동장치 안전모 착용](#)

9월 :

[메타버스](#)
[오징어게임](#)
[스포츠와 팬데믹](#)

10월 :

[K-로켓 누리호 발사](#)
[OTT 산업의 경쟁](#)

11월 :

[요소수 대란](#)
[신종 바이러스 오미크론](#)

12월 :

[백신패스 강화 & 워드 코로나 중단](#)
[부산 해상 도시 프로젝트](#)

1. [바이든 시대 & 미국 의회 습격사건](#)
2. [도쿄 올림픽 개막](#)
3. [K-로켓 누리호 발사](#)
4. 오미크론 정도..?

2022년 경제이슈(이성일)

-[wiki 2022 이슈 총정리](#)
-[BBC 국제이슈정리](#)
-[국내 경제 이슈 정리](#)

-[KDC분류](#)

-9월 경제 이슈

1월: LG에너지솔루션이 1월 27일 코스피에 상장했다

2월: 2월 24일 러시아의 우크라이나 침공으로 촉발된 러·우 전쟁이 **10개월** 넘게 이어지고 있다. 전쟁은 식량과 에너지, 원자재 가격을 끌어올리며 물가 폭등을 부추겼다.

2월 27일 러시아·우크라이나 전쟁 발발

3월: 윤석열 정부는 주식양도세 폐지와 공매도·물적분할 제도 개선을 골자로 한 자본시장 공약을 들고 **3월 9일** 출범했다.

4월: 금융당국이 **4월 20일** 뮤직카우의 음악 저작권료 참여 청구권을 투자계약증권으로 인정하면서 조각투자가 제도권 안으로 들어왔다.

4월 26일 일론 머스크가 트위터를 **55조원**에 인수하기로 결정하였다.

5월: 5월 김치코인(한국산 가상자산) 루나·테라 폭락 사태가 가상자산 시장을 덮쳤다.

6월: 6월, '尹 사단 막내' 이복현 금융감독원장 취임 **6개월**간 불법 공매도나 불공정거래에 엄단을 예고하며 추진력을 보여줬다.

6월 10일 - 미국의 5월 소비자물가지수(CPI)가 8.1%로 41년만에 최고를 기록하였다.

7월: 금융당국은 불법공매도 엄벌 의지와 함께 공매도 제도 개선책을 내놔다.

8월: 8월 8일, 80년 만의 기록적인 폭우로 한국투자증권의 **HTS, MTS**가 먹통이 됐다.

9월: 9월, 코스피 25% 하락...환율 13년 6개월 來 최고치

9월 13일 - 제74회 프라임타임 에미상 시상식에서 영화감독 황동혁과 배우 이정재가 드라마 《오징어 게임》을 통해 비영어권 드라마 그리고 아시아인 최초로 드라마부문 각각 감독상과 남우주연상을 수 상하였다.

10월: 레고랜드 사태는 강원도가 빚보증 의무 이행을 거부하면서 불거졌다.

10월 15일 - 판교 SK C&C 데이터센터에서 화재가 발생해 카카오톡 및 카카오에 기반한 모든 서비스가 마비되는 사태가 발생하였다.

10월 27일 - 이재용이 삼성전자 회장으로 취임했다.

11월: 11월 10일 한국거래소의 **67년** 독점 체제를 깬 대체거래소(ATS)가 출범을 예고했다.

12월: 12월 25일 여야는 금융투자소득세(금투세) 시행을 **2년** 유예하기로 합의했다.

<이슈>

1. 이슈 선정의 기준이 모호하다
2. **사기꾼 모멘트**
3. 도서관 선정기준? 아니면 인기도서만 할 것인가?

4.

<분석 가이드>

1. 전월 대비 대출이 증가한 주제 확인 후 이슈와 매칭

ex)

전월 대비 **22년 7월**의 대출 인기상승 주제: 자연과학,증가량:**12**

전월 대비 **22년 8월**의 대출 인기상승 주제: 자연과학,증가량:**12**

전월 대비 **22년 9월**의 대출 인기상승 주제: **사회과학**,증가량:**19**

전월 대비 **22년 10월**의 대출 인기상승 주제: 문학,증가량:**2**

22년 8월에는 없는 **22년 9월**의 인기 대출 도서

임지 센스 :한 번의 선택으로 부의 계급을 높이는 **부동산** 투자의 감각

나는 **4시간만 일한다** :디지털 노마드 시대 완전히 새로운 삶의 방식

(박병창의 돈을 부르는) 매매의 기술

우리는 어디서 살아야 하는가 :인문학자가 직접 고른 살기 좋고 사기 좋은 **땅**

세븐 **=7년** 안에 경제적 자유를 만드는 **7가지 비밀 /Seven**

나는 **부동산**과 맞벌이한다 :배우자 대신 꼬박꼬박 월급을 가져오는 시스템 만들기

9월 핵심 이슈: **9월, 코스피 25% 하락...환율 13년 6개월 來 최고치 -> 원화가치 하락-> 부동산 자산에 주목**

2. 경제이슈의 종류를 나눠서 어떤 종류의 이슈가 나왔을때, 분류기준

3. 월간 비교 지표 : 1. 대출건수, 2.

중간 개선

- 매칭

<가설 = BIG QUESTION>

-
- 사회적 이슈 자체가 도서에 영향을 미칠 수 있는가?
- 어떤 종류의 이슈가 영향을 주로 미치는가?

19년

1. **7월** 한일 무역분쟁 - 성운
2. **12월 covid** 개막

20년

- 3.

21년

4. **1월** : 바이든 시대 & 미국 의회 습격사건
5. **7월** : 도쿄 올림픽 개막
6. **3월** : K-로켓 누리호 발사
7. **11월** : 오미크론 발생
8. **12월** : 위드코로나 중단

22년

9. 1월: LG에너지솔루션이 1월 27일 코스피에 상장했다
10. 5월: 5월 김치코인(한국산 가상자산) 루나·테라 폭락 사태가 가상자산 시장을 덮쳤다.
11. 9월: 9월, 코스피 25% 하락...환율 13년 6개월 來 최고치

• 320 경제학

- 321 경제학론
- 322 경제정책
- 323 산업경제 일반
- 324 기업경제
- 325 경영
- 326 상업·교통·통신(공익 사업)
- 327 금융
- 328 보험
- 329 재정

전처리 코드

1. 파일로드
2. 매달 주제별 도서수
3. 전월 대비 대출이 증가한 주제
4. 전월에 없는 해당 달에 증가한 도서 제목 출력

<썬 막>

1. 각 년도별 월별 데이터 로드 (월별, all 전체, 1000개) - clear
2. 주제별 대출건수 합 월별 비교(증가량 파악) - 일요일 전까지
3. 2번 결과와 이슈와 얹은 연결 - 일요일

(시각화 코드는 통일)

ex) “인기대출도서_2022_01.csv” 추천..

최종 -> 이슈량 결과량 매칭 후 연관 보완

숙제 -> 더해서 전월과 빼고 특이점 - 이슈 찾아보기

아껴돌 아이템 (순위별 가중치)

회의록 - 강영민 2022.05.18

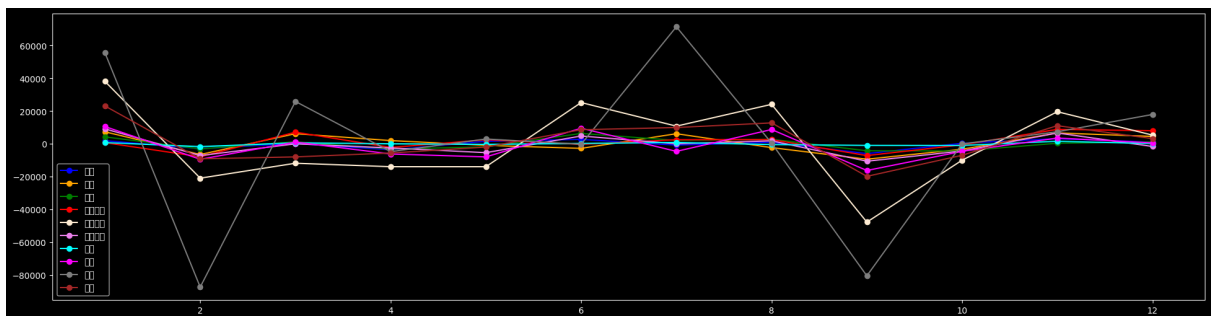
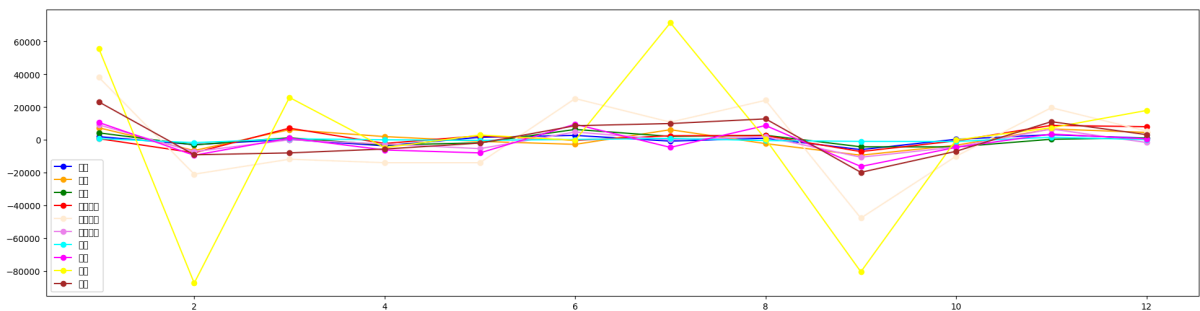
2019년 대출증가량 분석 - 김성윤

19년도 1월의 인기상승 주제 : 문학, 대출 증가량 : 55470
19년도 2월의 인기상승 주제 : 예술, 대출 증가량 : -1685
19년도 3월의 인기상승 주제 : 문학, 대출 증가량 : 25859
19년도 4월의 인기상승 주제 : 철학, 대출 증가량 : 2003
19년도 5월의 인기상승 주제 : 문학, 대출 증가량 : 2978
19년도 6월의 인기상승 주제 : 자연과학, 대출 증가량 : 25187
19년도 7월의 인기상승 주제 : 문학, 대출 증가량 : 71391
19년도 8월의 인기상승 주제 : 자연과학, 대출 증가량 : 24104
19년도 9월의 인기상승 주제 : 예술, 대출 증가량 : -916
19년도 10월의 인기상승 주제 : 총류, 대출 증가량 : 343
19년도 11월의 인기상승 주제 : 자연과학, 대출 증가량 : 19605
19년도 12월의 인기상승 주제 : 문학, 대출 증가량 : 17926

2월 이슈 : 북미 정상회담, 평창 올림픽

7월 이슈 : 메이저리그 올스타전, 에볼라 유행, 금리 인하, 호날두 노쇼, 일본의 공격

9월 이슈 : 태풍 링링, 한국의 일본 수출 규제



(한글 폰트 이슈)

- => 문학은 등락이 굉장히 심함
- => 다른 종류 다 증가할 때 감소하는 종류 찾기
- => 전반적으로 독서량이 하락하는 시즌, 2월, 9월
- => 전반적으로 증가한 12월, 문학만 높은 7월
- => 자연과학도 등락의 변화도 심했음
- => 방학 이슈라고 하기에는 2월이 최악을 기록하고 다시 3월에 증가함.
- => 다양한 이슈가 존재할 수 있다고 판단.

=> 모델을 통해서 분석해보는 것도 좋아보임 => 아니면 상관계수 정도 파악해보기
=> 아니면, 상위 인기대출 도서들의 특징 파악(책 이름, 저자, 출판사, 출판년도 등 다양한 변수들의 상관관계를 생각해보고 순위 자체의 특징도 다양하게 바라볼 필요가 있어보임)
=> 이런 주기성이 발생하는 것 역시 발견이고 분석의 이유를 찾아보는 것도 좋아보임
=> 4년동안의 분류별 주기성을 발견하고 그 이유를 분석(주기성을 입증하기 위해서 더 많은 년도를 찾아봐도 좋음) + 특정 해의 이상치가 발생한다면 그것에 대해서 이유를 분석
=> 영민님이 말씀하신 것처럼 서점 데이터를 살펴봐서 비교해보는 것도 좋아보임
=>
=> 주제 : 도서관 대출정보를 통해서 새로운 인사이트를 제공

9월에 상당한 메타버스관련 이슈가 있었음 -> KDC 300번대의 대출 증가량이 가장 컸음
2020년 대출 증가량 분석 - 이호건

20년도 1월의 인기상승 주제 : 문학, 대출 증가량 : 219204
20년도 2월의 인기상승 주제 : 예술, 대출 증가량 : -2559
20년도 3월의 인기상승 주제 : 예술, 대출 증가량 : -1099
20년도 4월의 인기상승 주제 : 문학, 대출 증가량 : 86413
20년도 5월의 인기상승 주제 : 문학, 대출 증가량 : 84842
20년도 6월의 인기상승 주제 : 예술, 대출 증가량 : -82
20년도 7월의 인기상승 주제 : 문학, 대출 증가량 : 118811
20년도 8월의 인기상승 주제 : 자연과학, 대출 증가량 : 4725
20년도 9월의 인기상승 주제 : 예술, 대출 증가량 : -731
20년도 10월의 인기상승 주제 : 문학, 대출 증가량 : 81849
20년도 11월의 인기상승 주제 : 문학, 대출 증가량 : 28320
20년도 12월의 인기상승 주제 : 예술, 대출 증가량 : 347

11월에 오미크론 관련 이슈가 있었음 -> KDC 300번대의 대출 증가량이 가장 컸음

전월대비 01의 가장 큰 증가폭을 보인 KDC는 800번대이며, 증가량은 80257건입니다.
전월대비 02의 가장 큰 증가폭을 보인 KDC는 200번대이며, 증가량은 4049건입니다.
전월대비 03의 가장 큰 증가폭을 보인 KDC는 800번대이며, 증가량은 35708건입니다.
전월대비 04의 가장 큰 증가폭을 보인 KDC는 800번대이며, 증가량은 3037건입니다.
전월대비 05의 가장 큰 증가폭을 보인 KDC는 400번대이며, 증가량은 5675건입니다.
전월대비 06의 가장 큰 증가폭을 보인 KDC는 800번대이며, 증가량은 12396건입니다.
전월대비 07의 가장 큰 증가폭을 보인 KDC는 800번대이며, 증가량은 93737건입니다.
전월대비 08의 가장 큰 증가폭을 보인 KDC는 400번대이며, 증가량은 5670건입니다.
전월대비 09의 가장 큰 증가폭을 보인 KDC는 300번대이며, 증가량은 1979건입니다.
전월대비 10의 가장 큰 증가폭을 보인 KDC는 700번대이며, 증가량은 1218건입니다.
전월대비 11의 가장 큰 증가폭을 보인 KDC는 300번대이며, 증가량은 3223건입니다.
전월대비 12의 가장 큰 증가폭을 보인 KDC는 800번대이며, 증가량은 21564건입니다.

2022년 대출증가량 분석 - 이성일
1,5월 발생한 경제 이슈에 대응되는
주제는 없었음 (9월에 경우 저널리즘 등이 있는 총류 섹션이 증가하긴 했음)

전월 대비 22년 1월에 가장 큰 대출증가폭을 보인 주제: 문학, 증가량:60884

전월 대비 22년 2월에 가장 큰 대출증가폭을 보인 주제: 예술, 증가량:-829

전월 대비 22년 3월에 가장 큰 대출증가폭을 보인 주제: 문학, 증가량:45529

전월 대비 22년 4월에 가장 큰 대출증가폭을 보인 주제: 종교, 증가량:1200

전월 대비 22년 5월에 가장 큰 대출증가폭을 보인 주제: 예술, 증가량:1821

전월 대비 22년 6월에 가장 큰 대출증가폭을 보인 주제: 기술과학, 증가량:2035

전월 대비 22년 7월에 가장 큰 대출증가폭을 보인 주제: 문학, 증가량:74788

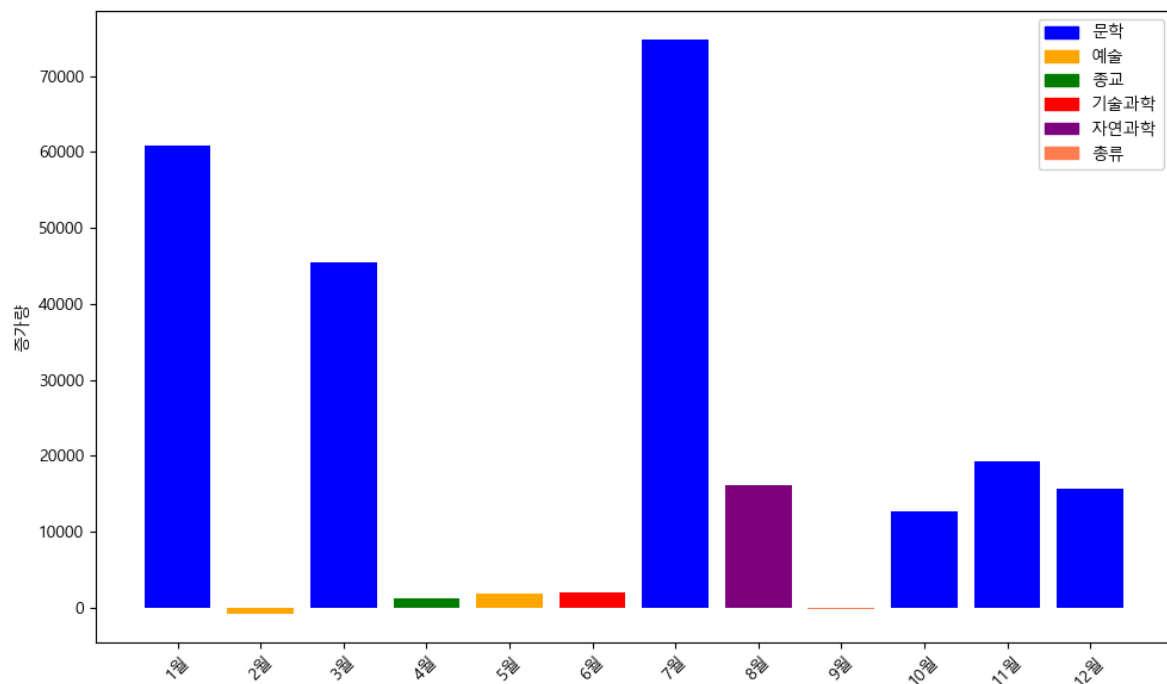
전월 대비 22년 8월에 가장 큰 대출증가폭을 보인 주제: 자연과학, 증가량:16215

전월 대비 22년 9월에 가장 큰 대출증가폭을 보인 주제: 총류, 증가량:-152

전월 대비 22년 10월에 가장 큰 대출증가폭을 보인 주제: 문학, 증가량:12712

전월 대비 22년 11월에 가장 큰 대출증가폭을 보인 주제: 문학, 증가량:19271

전월 대비 22년 12월에 가장 큰 대출증가폭을 보인 주제: 문학, 증가량:15616



중간발표 ppt 피드백 김성윤

- > 중간 발표 가이드와 다른 점이 상당 수 존재해 보임. (배경, 목표, 데이터, 기대효과)
- > ppt에 기대효과가 존재하지만, 단순 질문의 꼬리만 연결되는 느낌
- > ppt가 상당히 산만하고 정리가 되어 있지 않은 느낌을 받았음.
- > 지난 주제에 대한 언급은 굳이 시간을 투자하지 않아도 좋아보임. (그냥 새롭게 시작하는 느낌으로 제시)
- > 배경과 목표, 기대효과에 대해서 모호해 보여 정해봤음.

=> 알아엮어야 할 것 알아보임.

배경 : 도서관 데이터를 찾게되고 가설을 세우게 됨. 도서관 데이터를 분석한다 -> 도서 경향을 알 수 있다. -> 현대인들의 생각(니즈)를 유추할 수 있다. -> 다양한 사업적 가치로 활용할 수 있다.

목표 : 도서관 데이터를 통해서 도서 경향을 파악하고 영향 요인들을 파악하기

기대효과 : 도서 경향을 기반으로 한 도서관 프로그램 컨설팅, 영향 요인 파악을 통한 도서 경향 예상, 도서 경향을 기반으로 한 현대인들의 심리 파악

데이터 : '도서관 정보나루-인기대출도서' 데이터 이용

=> 컬럼 소개

=> 사용할 데이터들의 정보(기간, 장르 등등)

일정 계획 및 역할 분담 계획 대비 진행 상황

<역할 분담>

1. 데이터 선정 -> **clear**

2. 데이터 수집과 전처리 -> **clear**

성윤 : 19년도 인기대출도서 데이터 수집 및 전처리

영민 : 20년도 인기대출도서 데이터 수집 및 전처리

호건 : 21년도 인기대출도서 데이터 수집 및 전처리

성일 : 22년도 인기대출도서 데이터 수집 및 전처리

3. 데이터 분석 -> 기본적인 분석 기법을 통한 분석 **clear**

모든 인원 : 데이터 분석 기법은 선정

성윤 : 19년도 인기대출도서 데이터 분석

영민 : 20년도 인기대출도서 데이터 분석

호건 : 21년도 인기대출도서 데이터 분석

성일 : 22년도 인기대출도서 데이터 분석

4. 결과 해석-> 기본적인 결과 해석 **clear**

성윤 : 19년도 사회적 이슈 조사 및 연관성 파악

영민 : 20년도 사회적 이슈 조사 및 연관성 파악

호건 : 21년도 사회적 이슈 조사 및 연관성 파악

성일 : 22년도 사회적 이슈 조사 및 연관성 파악

<일정 계획>

중간 발표 전 : 데이터 선정, 간단한 데이터 분석을 통해서 결과 초안 제시 -> **clear**

중간 발표 후 : 다양한 컬럼과 데이터를 활용하여, 도서 경향과 영향 요인 파악

분석 및 실험 결과와 해석

분석 내용 : 19~22년도별 월별 대출 증가량 파악

실험 결과 : 특이점 발견

해석 : 영향 요인 파악

앞으로 진행할 분석 내용: 가설(hypothesis)을 포함

1. 도서 경향의 특이점과 사회적 이슈 사이의 상관관계가 있을 것이다.
2. 서점 데이터를 이용한다면 영향 요인에 대해 더 인과성을 설명할 수 있을 것이다.
3. 더 많은 년도를 분석한다면 **general**한 도서 패턴을 만들 수 있을 것이다.
4. 다양한 컬럼을 활용하여 분석한다면 **general**한 도서 패턴을 만들 수 있을 것이다.

<피드백>

분석결과를 보여줄 땐 좀 더 도구의 결과

-> 시각화를 좀 더 해달라

-> 현재 텍스트가 너무 많음

앞으로 할 것

-> 결국은 지금 가지고 있는 데이터가 도서 이기에 추가적인...

책에 대한 리뷰 + **yes24**, 알라딘, 책에 대한서평 서문

-> 그걸 통해서 어떤 분야 -> 어떤 책인지 특징을 고도화하도록 활용하기
를 한다면 재밌는 요소가 나올 수 있을 것 같다

-> 사회적이슈에 대해 더 수집하지 말아라(오래걸릴 수 있다, 보충 설명 자료로 활용하기)

5/23 AM12:30 회의

방향성 구체화

1. ISBN 활용하기

2.독서트렌드 -> 장르->KDC->서점 데이터는 kdc 분류가 없다

3.리뷰활용하기 -> 전문가 서평 분석 후 긍정적인 서평과 부정적인 서평이 대출에

미치는

영향

4. 서점 데이터와 도서관 대출 데이터의 비교 (서점 데이터는 알라딘으로)

5. wow point : 일치한 파트에 대해서 서점 리뷰데이터 크롤링

6. (리뷰데이터 - 일치한 파트) => 몰타기된 서적들 => 실제 리뷰도 좋은가?

7. 저희는 도서 단위로 분석?

줄거리,

인기서적 줄거리에 어떤 형용사가 들어있나

-> 도서관, 서적(줄거리) 고점 베스트셀러들의 키워드? 이슈가 해당 달에 터졌는가? 해당
키워드와 관련된

=> 콘텐츠(줄거리), 작가, 출판사,

기사(내용) 크롤링 - 서점 데이터(줄거리) 크롤링 => 교수님 거봐 안된다 했잖아! 시간
없다했잖아, 오? 이걸 해내내 연구실 들어올 생각/....

시각화 부족 현상

=> 기분이 상당히 좋습니다

=> 크롤링 기술 학습

=> 금요일까지 크롤링 데이터화하기

=> 좋습니다



목요일까지

=> 베스트 셀러 데이터 4년치

=> 서점 데이터 전처리, 크롤링 공부

=> 오후 3시

<5/25 회의>

이슈

1. 정제되지 않은 데이터
2. 판매량을 유추할 지표 -> Big Issue -> 순위 증가량, 1000위 이내의 주제별 순위 값
3. 구글 기사 크롤링 이슈

2번 이슈 해결

1. 각 도서별 순위 증가량
2. 1000위 이내의 주제별 순위 합
3. 1000위 이내의 주제별 개수 증가량
4. 도서관 데이터

-> 주제 -> 도서의 제목 -> 키워드 -> 기사와 연결 -> 인사이트

5. 저자, 출판사

도서관 데이터 - 이슈에 민감하게 반응하지 않음

[가설] 서점 데이터는 이슈에 민감하게 반응할 것이다.

서점 데이터 분석 -> 도서관 데이터와 대조

특정 분야들은 양측 모두 이슈에 민감하게 반응

특정 분야들은 한쪽만, 혹은 모두 이슈에 반응하지 않음

왜 그런지 분석 & 줄거리도 필요하면 ㄱㄱ

<최종까지 해야할 일>

도서관 데이터 분석

서점 데이터 분석

줄거리 자연어 처리

한줄평 자연어 처리

한줄평 크롤링

기사 자연어 처리

기사 크롤링

데이터 총합

인사이트 제시

서점 책과 도서관 책의 공통점 -> 왜 찾아야 하는가?

=> 진짜 읽을만한가? 물타기된 책이 아닌가? 어떻게 증명가능한가?

=> 도서관에 비치된 책들의 공통점?은 무엇일까? 교양? 인생에 도움되는?=> 도서관에 비치되면서 인기 급상승한 도서는 무엇인가?

할 일 정리

1. 도서관 데이터 1000권 내 분류별 도서 수 전처리(서점 데이터와 같은 방식)
2. 서점 데이터도 전처리 진행
3. 도서관 데이터 - 서점 데이터 fusion -> 1인1역 : 시각화 & 주제 뽑아내기
4. 독특한 특정 주제 선정
5. 그 주제의 도서의 제목(+줄거리)을 분석 -> 1인1역 : 자연어처리 모델
6. 키워드 -> 1인 1역 : 시각화(호건, 정병1)
7. 크롤링한 기사 내용 분석
8. 키워드끼리 매칭 비율정도(accuracy) 파악
9. accuracy로 관계성 설명

다음 회의 : 5월 30일(화) - 15시

그때까지 뭐해야와야돼요 ..????

1. 데이터 전처리
2. fusion을 통해서 월별 주제를 뽑아냄 -> 도서관-서점 상관관계 파악
3. 주제별 도서 제목 분석해서 키워드 뽑아내는 자연어처리 모델 -> 89% clear (very good!!)
4. 기사 스크랩 코드
5. 스크랩 내용 분석해서 키워드 뽑아내는 자연어처리 모델
6. 키워드간 유사도를 점수로 환산 -> 관계성이 어느정도 존재하는지 제시 (GG)

if (계획이 순탄)

비대면

else

대면

5/30

가정 : 이슈에 의해 도서 트렌드가 변경되었다.

도서관, 서점 데이터 비교 같이 증가한 주제 분석

-> 흐름이 있다 -> 트렌드가 있다

서점과 도서관이 비슷한 책이 있는가?

관련 서적에 대한 리뷰를 본다 -> 긍정적인가? 부정적인가? 키워드를 통해 분석
긍정적이다? 읽을만하다 부정적이다? 그저 트렌드 때문에 뜬 의미없는 책이다.

주제를 완전 같아 쏘여야 함.

-> 지금은 뭔가 보완으로 가야하지 같아쏘는다면 가벼운 분석까지만,, 가능하지 않을까
걱정이 됩니다.

<지금 방향성에 대해서 걱정*불편하신 점>

- 전처리 방식 (제일 높은 분포를 비교한다. => 도서관과 서점의 겹치는 점이 문학
뿐...)
- 도서관과 서점 데이터를 유의미하게 엮지 못하겠다! (마땅한 연결고리가 보이지
않는다)
- 깔삼하게 떨어지지 못한다!
- 항상 일반적인 개형을 띄는 문학에 계속 막힌다. => 테마를 하나(사회과학)에만
집중해보자
- 일단 문학과 총류를 삭제해야합니다!!

데이터의 특징이 너무 무난하다!!!

“둘의 상관관계가 없다.”

두 가지 해결방안

- 다른 대안을 찾는다. -> 머리를 맞달고 다시 데이터를 어떻게 이용해 먹을까...
고민,,, 순탄하지 않는다. 의견을 모으기 쉽지 않다. 목요일에 미팅을 해서 간단한
브리핑을 해야하는데... 새로운 모색점을 찾고 있다. -> 저희는 F입니다.
- 상관관계가 없는 것을 이용한다. -> 대수롭지 않게, 상관관계 왜 없을까? 혹은
상관관계를 분석하지 않고 각자의 데이터로서 이용하기. 이슈를 찾으면,
서점데이터와 비교했고, 도서관이랑도 비교했다. 이렇게 했을 때, 두 데이터 중에
이슈와 더 관련 있던 건 어느 데이터였다.
상관관계를 어떻게 파악했나 수치적으로,,

이슈와 관련있어보이는게 저희 사회과학... 사회과학 위주로 봤다. => 사회과학 크게 변화가 없다. => 실제로 이슈와 도서는 크게 연관성이 없었다.(이슈가 있으면 도서 증가한다.(통념) -> 의미가 있다.)

정확하게 둘다 있습니다. 저희 시도를 절대 묵인하지 않고, 무시하시지도 않는다. 하지만, 거기에 대해서 어떻게 해결나갔는지, 결국에는 끝은 있어야합니다. 성적에 해가 되지 않는다.

<벽을 만났다. 해결했다.>

1. 데이터는 사회과학으로 축소합니다.
2. 사회과학의 변화 추이를 살핍니다. (도서관, 서점 데이터 살펴도 되고, 하나만 해도 됩니다. 둘다 파악하고 선택하면 됩니다.) + (서점, 도서관 추이 파악)
3. 팍팍 뜨는 시점에 이슈를 찾아봅니다. (서점에 비해 도서관이 밀리는 패턴이 있다면, 서점의 팍팍 시점을 기준으로)
4. 이슈들의 키워드를 파악(기사 스크랩?)하고 사회과학 증가할 때, 상위 1000권 내에 존재하는 도서들의 특징
5. 두 키워드 리스트 간 상관관계 혹은 점수 파악

좋은 팀플 남길 간절히 바랍니다... 호건님, 성일님, 영민님 합작입니다

Issue 1)

(제목, 줄거리, 리뷰)

리뷰 좀 더럽긴 할 거 같다 -> 힘들다

제목도 힘들다.예) 역행자,1q84 ->?

줄거리 => OK (1. 크롤링, 2. 다른 서점데이터 줄거리 isbm 일치 이용)

<https://www.aladin.co.kr/shop/wproduct.aspx?ItemId=309295168>

(좌 서점, 우 도서관)

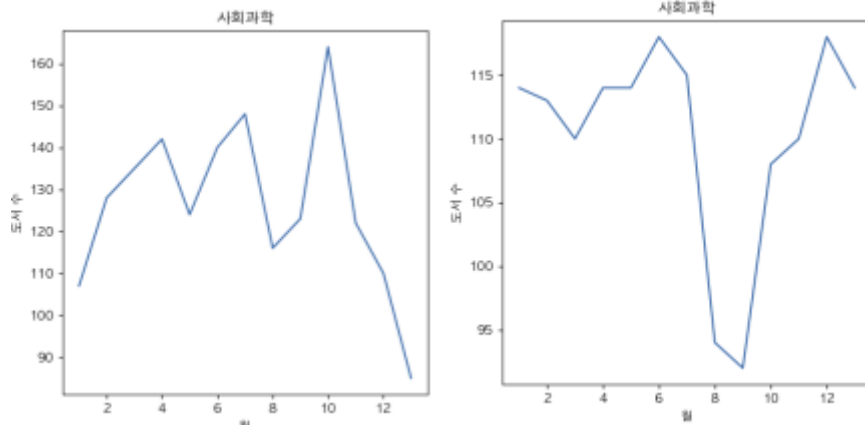
Issue 2)

서점 베스트 셀러가 더럽다. -> 현실남매, 원피스 “부동의 상위권 총”

<5/30>회의 정리

1. 데이터 수집(네이버 기사 크롤링, yes24 데이터와 매칭)
2. 모델 고도화 (선택사항)
3. 도서관, 서점 데이터 처리 => 주제만 뽑아내자
- 4.

제가 죄송합니다. 저도 완벽을 추구하고 싶지만, 그럴 시기가 아닙니다. 너무 상처받지 않으시길...



데이터 수집, 데이터 병합, 데이터 전처리, NLP 키워드 추출, 크롤링, 상관관계 분석
+ 결과...

계획에 없던 벽을 만났다.

두 가지 해결방안

- 다시 뒤로 돌아간다. 다른 길을 찾기
- 벽을 넘어서는 겁니다.
- 벽을 뚫기 압!! 빠세~!!

데이터를 어렵게 (크롤링), 성능 높이려 했다. -> 한심해보일 수 있다.

금부정 기준 -> 모델 성능, 분석을 하려는건지. 금부정 판단 이후에 => 읽을만했다. 트렌드에
뜬 허점이다.

NLP -> 감성분석, 분류 문제 => 감성분석 네이버 영화 리뷰 데이터 -> 긍정적, 부정적 예측!

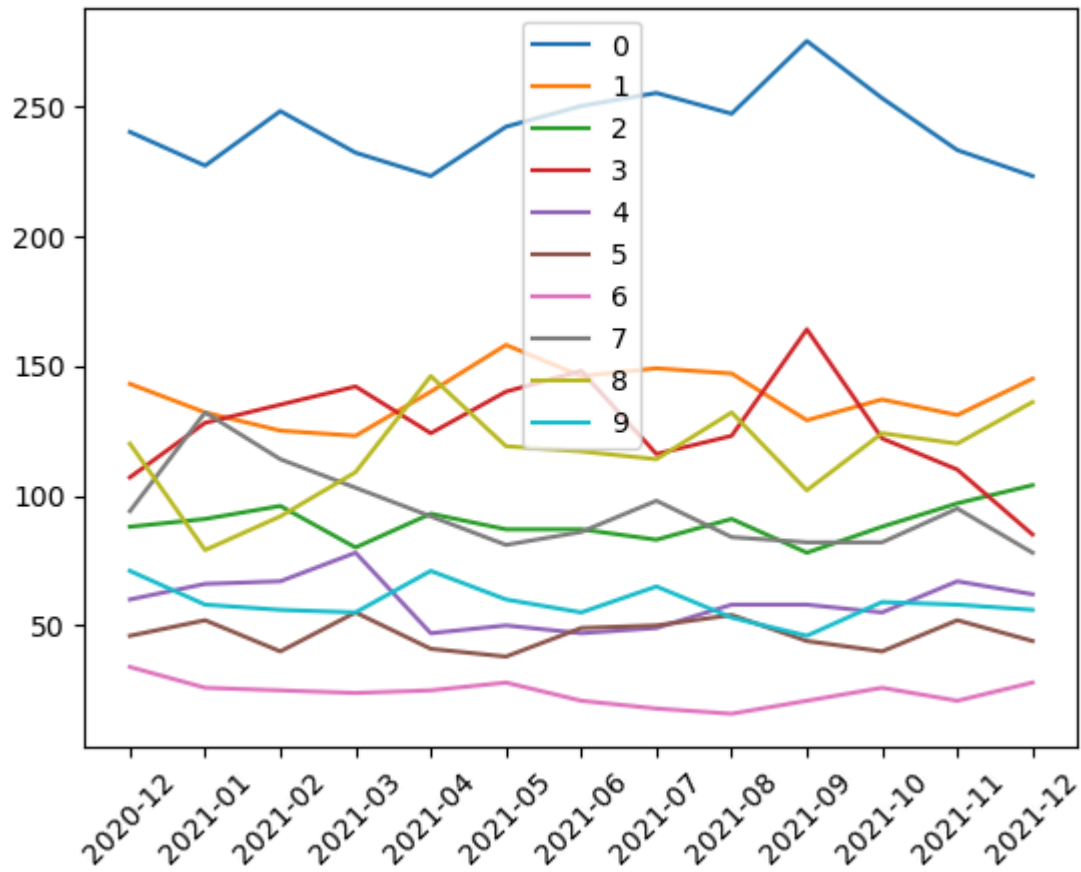
성능높이기

스팸 분석 -> 예측 성능을 높이기

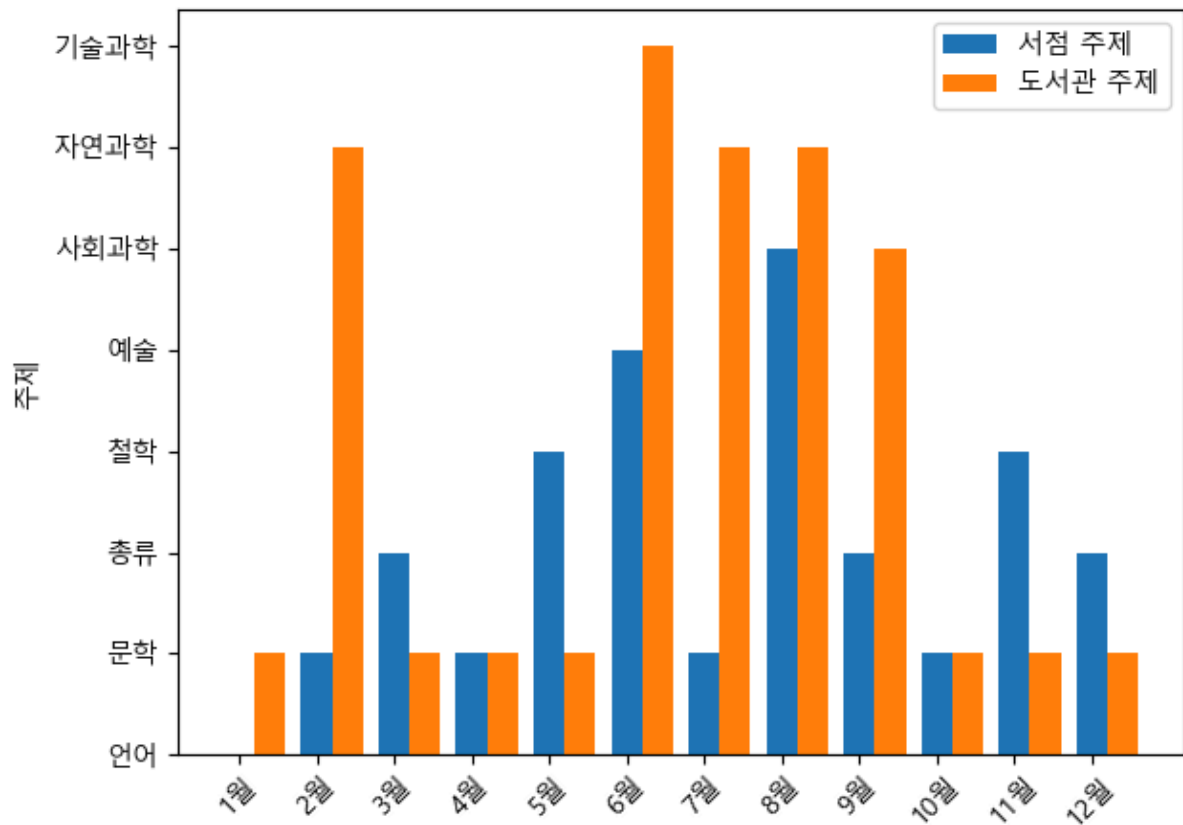
부의 추월차선

도서관, 서점(서점에서 부정적 긍정적 리뷰 보면 되는데 도서관 왜 엮어야 하는가?)

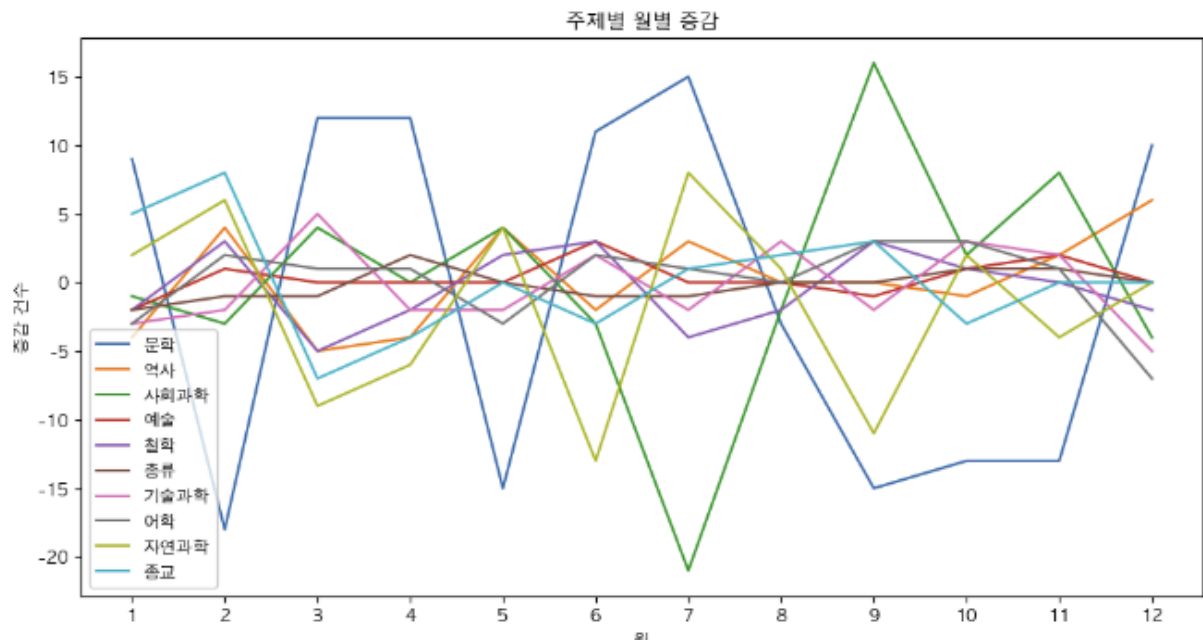
회의록 - 김성운



월별 서점과 도서관의 주제



'0': '종류',
'1': '철학',
'2': '종교',
'3': '사회과학',
'4': '자연과학',
'5': '기술과학',
'6': '예술',
'7': '어학',
'8': '문학',
'9': '역사'



<도서관 이용률 관련 이슈 정리>

- 모든 정보를 유튜브로 혹은 검색엔진을 통해서 얻는다. => 독서는 어느 영역에 존재해야 하는가...?
- 갈수록 하락세...
- 이용률 감소 원인 : 도서관까지 너무 멀다, 한줄평 - 온라인 시대로의 세대교체.?
- 종이 책 != 정보 습득의 수단,
- 다른 목적이 필요해 보임... 이용률을 높이기 위해서는 종이 책만을 원하는 사람들을 집중 타겟. 매니아 층...? 고연령대...

결국 '도서 대출 수나 구매 수가 높아지는 시점을 파악 혹은 예상해야한다!' 라는 결론이 나와야 함.

1. 사회적 이슈 2. 서점의 특정 주제의 구매 수 3. 도서관의 특정 주제의 대출 수 세 개의 상관관계를 파악하는 중.

기대효과 : 도서관 측에 제시해서 ~~~하도록 한다.

<Do List>

- 스토리 라인 (디지털화로 인해 적자를 최소화하기 위해서 발매에 신중을 가해야하는 출판사의 story)
- 키워드 이쁘게 뽑아내기 (konlpy를 통해서 키워드(tf-idf 점수 이용) 개선, but 완벽하지 않음을 줄거리로 보완 예정)
- 코드 자동화 및 통합
- 시각화

도서 데이터

특정 주제의 도서 데이터

특정 주제의 도서 데이터의 제목

특정 주제의 도서 데이터들의 키워드

특정 주제의 도서 데이터들의 키워드와 관련된 뉴스 기사 수

<발표 흐름>

1. 디지털 시대에 자연스럽게 스며든 청중들을 인식
2. 종이책을 출판하는 출판사와 작가의 입장에서 생각을 해보자 (변화를 기준으로)
3. 사업을 활성화하기도 접기도 애매한 상태임을 강조
4. 최선이 적자를 줄이는 것임을 강조
5. 그러기 위해서는 가장 핫할 시기에 책을 출판하는게 핵심
6. 그렇다면 (핫할 시기 = 사람들이 그 책의 주제를 원하는 시기)를 파악하는 것이 필요합니다.
7. 그래서 준비했습니다. 우리의 분석

8. 도서관 데이터와 서점 데이터를 분석하며 도서 패턴을 파악하고 영향 요인을 파악하기 위해 가설(사회적 이슈가 영향을 미친다)을 세우고 입증하려는 시도를 함
9. 결과적으로 ~~함을 알 수 있었다.
10. 이를 통해 우리는 어떻게 할 수 있다는 것을 알 수 있다.
11. 앞으로 어떤 과정을 통해서 조심스럽게 앞으로 그려질 미래를 예측해봄(예, 출판사의 적자 최소화가 아닌 오히려 이윤이 발생하는 것. 단순 추측이 아닌 근거 자료 찾아보기, 책이 대박나면 출판사가 이윤을 본다는 것. 출판사 계의 김은숙 작가님을 만들자!!)

<필요한 팩트> (-> 데이터 시각화와 연결됨)

팩트 1) 도서관 데이터와 서점 데이터를 통한 도서 패턴 결과

팩트 2) 가설에 대한 결과

TF-IDF 점수 이용 키워드 추출

1. 부의: 1.0
2. 부자: 0.93
3. 부동산: 0.75
4. 법칙: 0.62
5. 허영: 0.57
6. 시작: 0.51
7. 재테크: 0.48
8. 가지: 0.48
9. 트렌드: 0.46
10. 투자: 0.46

1. 부의: 100%
2. 부자: 93%
3. 부동산: 75%
4. 법칙: 62%
5. 허영: 56%
6. 시작: 51%
7. 재테크: 48%
8. 가지: 48%
9. 트렌드: 46%
10. 투자: 46%

이 %가 변하는 정도로 월별 이 단어의 중요도가 변한 것을 알 수 있습니다.

즉 이 단어가 갑자기 중요해졌다는 지점을 알 수 있게 됩니다

+ 특정 %를 threshold로 정해서 그 수치 이상은 키워드로 판단했다라고 주장

담아야 하는 내용:

- 배경과 목표(어떤 문제를 타겟하는지, 그 문제가 왜 중요한지),
- 가설이 무엇인지 (있다면), 어떤 방법을 사용하는 지,
- 결과가 무엇인지, 결과의 의미는 무엇인지,
- 한계점과 가능한 향후 방향, 팀원별 기여도 및 기여 내용 등에 대해 발표

교수자는 각 팀의 프로젝트 진행 히스토리를 알고 있지만, 모른다고 생각하고 평가할 것입니다.

청중을 고려한 발표를 준비하기 바랍니다. 각 팀이 고생한 것은 잘 알고 있으나, 얼마나 고생했는지 어떤 분석을 했는지 등을 나열하는 것은 좋은 발표가 아니며, 높은 점수를 받지 못할 것입니다.

팀원 간 상호 평가에 의해 우수 참여자, 참여 저조자를 결정합니다. 우수 참여자는 + 학점을 받습니다.

(분반별) 교수자 평가 1등 팀, 상호평가 1등 팀, 팀워크 1등 팀은 전원 + 학점을 부여합니다(참여 저조자 제외). 팀워크 1등 팀 판단 기준은 프로젝트 발표를 마친 후 공개합니다.

성일님 :

배경: 영상매체등으로 인한 판서매체의 약세 -> 해결방안: 도서관 이용 활성화 -> 방법: 독서 트렌드에 영향을 주는 요인 분석 -> 이슈에 따라 ? -> 이슈 관련된 책 증가시키기

이슈와 직접연결될거라고 생각하고 키워드를 대입했으나 그렇지 않았다 또한 원하는 책이 없어서 도서관을 이용 안하진 않는다.

도서관 이용하지 않는 이유			
시간이 없어서	주변에 도서관이 없어서	도서관 이용 필요성을 느끼지 못해서	내가 원하는 책이 없어서
☹ ☹ ☹ ☹	☹ ☹ ☹ ☹	☹ ☹ ☹ ☹	☹ ☹ ☹ ☹
35.8	10.1	51.7	1.2
39.5	8.9	49.6	1.0
31.9	11.3	53.8	1.4
56.7	20.3	19.7	2.2
47.5	17.9	31.8	1.2
44.4	17.2	34.3	2.2
46.1	10.4	41.5	1.5
37.8	5.5	55.1	1.4
9.9	4.5	83.6	0.0
53.8	17.3	23.1	3.8
60.6	24.2	15.2	0.0
47.2	20.8	29.2	0.0

성윤님 : 성일님 의견 피드백

조심스럽게 피드백을 하자면, 저도 도서관 활성화를 가장 먼저 떠올렸었습니다. 하지만, 한가지 의문이 '도서관 이용률 감소'에 가장 큰 영향을 미치는게 사람들이 좋아하는 책이 적어서인가?라는 생각이 들어서 짚짚했었습니다.

'이걸 왜 분석해야하나' 즉, 저희의 과정이 필요한 활동이라면 결과는 크게 중요하지 않을 수 있다고 생각합니다.

성윤님 : 디지털 시대로 가고 무료 콘텐츠가 많아져서 출판사의 매출이 감소하는 추세라고 합니다. 그러니까 출판사는 아무 시기에 책을 내면 안되고 최대한 인기가 있을 시기를 찾아서 발매? 해야합니다.

관련 기사 :

<https://www.readersnews.com/news/articleView.html?idxno=108673>

<https://appmax.tistory.com/11>

<https://www.giikorea.co.kr/report/infi1076888-global-publishing-market.html>

=> 앞으로의 전망이 좋은 출판시장에서 영업이익을 더 잘 챙기기 위해서 우리의 분석과 같은 작업이 필요하다

영민 :

왜 우리는 키워드를 통해 기사 수를 분석하는가?

기사 수가 많다는 뜻은 해당 키워드(주제)가 인기있었다를 증명 가능하기 때문이다.

이는 해당 키워드가 들어간 책이 인기가 많았다는 것 또한 증명이 된다.

우리의 분석 주제는 고로 시기와 직접적인 연관이 있게된다. 시기와 연관 지을 수 있는 것에 포커스를 맞춰보자면

- 책의 발매시기
- 도서시장의 흐름 파악 가능
- 이슈에 민감한 도서들이 무엇이 있나 알 수 있고 도서관은 해당 도서들을 굳이 비치할 필요가 없고 예산을 다른 곳에 효율적으로 사용할 수 있게할 수 있다. 여기서 다른 곳이라 함은 도서관에 사람들을 끌어모을 수 있는 방안을 모색하는 일을 말한다. 예를 들면 여유로운 예산으로 보다 일찍 냉난방을 시작하는 도서관, 도서관에서 주최하는 상금이 걸린 문학창작대회 등..

후에 나아가서는 단순히 경제뿐만 아니라 다른 주제의 책과 이슈를 분석해보며 이슈가 터지고 해당 이슈에 대한 책의 반응성이 언제 커지는지 분석해볼 수 있고 이는 출판사의 상업적 이용 가능성과 직접적인 연관이 있다.

흐름 변경안 추천(의견참고만..)

1. 디지털 시대에 자연스럽게 스며든 청중들을 인식
(디지털 시대도 좋지만 휘발성이 강한 중독성 강한 콘텐츠, 유튜브, 틱톡 쇼츠 릴스 점점 자극적이고 단기간에 쾌락을 얻는 콘텐츠에만 집중하는 실태)
2. 종이책을 출판하는 출판사와 작가의 입장에서 생각을 해보자 변화를 기준으로
(종이책보다는 그냥 출판사와 작가의 입장으로 하는게 어떨지.. 도서 이용률을 높이기 위해)
3. 사업을 활성화하기도 접기도 애매한 상태임을 강조
4. 최선이 적자를 줄이는 것임을 강조
5. 그러기 위해서는 가장 핫할 시기에 책을 출판하는게 핵심
6. 그렇다면 (핫할 시기 = 사람들이 그 책의 주제를 원하는 시기)를 파악하는 것이 필요합니다.
(또한 그 주제에 대한 명확한 키워드를 끼워넣는것도 하나의 방법 -> 기존 분석 코드를 통해 이번엔 기사의 제목을 정제해도 좋을듯)
7. 그래서 준비했습니다. 우리의 분석
8. 도서관 데이터와 서점 데이터를 분석하며 도서 패턴을 파악하고 영향 요인을 파악하기 위해 가설(사회적 이슈가 영향을 미친다)을 세우고 입증하려는 시도를 함
9. 결과적으로 ~~함을 알 수 있었다.

10. 이를 통해 우리는 어떻게 할 수 있다는 것을 알 수 있다.
11. 앞으로 어떤 과정을 통해서 조심스럽게 앞으로 그려질 미래를 예측해봄(예, 출판사의 적자 최소화가 아닌 오히려 이윤이 발생하는 것. 단순 추측이 아닌 근거 자료 찾아보기, 책이 대박나면 출판사가 이윤을 본다는 것. 출판사 계의 김은숙 작가님을 만들자!!)
 - 도서관장의 흐름 파악 가능
 - 이슈에 민감한 도서들이 무엇이 있나 알 수 있고 도서관은 해당 도서들을 굳이 비치할 필요가 없고 예산을 다른 곳에 효율적으로 사용할 수 있게할 수 있다. 여기서 다른 곳이라 함은 도서관에 사람들을 끌어모을 수 있는 방안을 모색하는 일을 말한다. 예를 들면 여유로운 예산으로 보다 일찍 냉난방을 시작하는 도서관, 도서관에서 주최하는 상금이 걸린 문학창작대회 등..

<최종 발표 가이드>

• 평가 기준: 결과의 타당성, 의의, 전달력 등을 종합적으로 판단

• 발표시간: 팀별 최대 **10분**

• 발표자료 제출기한: **6.5(월) 정오**

- 분석 결과 재현 가능한 코드, 회의록을 함께 제출
- 개인별 팀원 평가표 제출
- 웹 상에 유사한 코드가 존재할 경우 치팅으로 판단

담아야 하는 내용:

- 배경과 목표(어떤 문제를 타겟하는지, 그 문제가 왜 중요한지),
- 가설이 무엇인지 (있다면), 어떤 방법을 사용하는 지,
- 결과가 무엇인지, 결과의 의미는 무엇인지,
- 한계점과 가능한 향후 방향, 팀원별 기여도 및 기여 내용 등에 대해 발표

교수자는 각 팀의 프로젝트 진행 히스토리를 알고 있지만, 모른다고 생각하고 평가할 것입니다.

청중을 고려한 발표를 준비하기 바랍니다. 각 팀이 고생한 것은 잘 알고 있으나, 얼마나 고생했는지 어떤 분석을 했는지 등을 나열하는 것은 좋은 발표가 아니며, 높은 점수를 받지 못할 것입니다.

팀원 간 상호 평가에 의해 우수 참여자, 참여 저조자를 결정합니다. 우수 참여자는 + 학점을 받습니다.

(분반별) 교수자 평가 1등 팀, 상호평가 1등 팀, 팀워크 1등 팀은 전원 + 학점을 부여합니다(참여 저조자 제외). 팀워크 1등 팀 판단 기준은 프로젝트 발표를 마친 후 공개합니다.

=====

- 데이터 넣어서 코드 돌려보기 -> 시각화(결과) 방향성 정하기
- 배경과 목표 정하기 (왜 중요한지에 대해서 찾기) ->clear

- 한계점과 향후 방향 정하기

<표면적인 한계점>

1. 이슈 외 대출량/판매량에 영향을 끼치는 다른 요인들이 있어 결과의 정확도가 문제요인이다.
- 2.

<기술적 한계점>

1. 한정적인 데이터로 분석을 하다보니 다양한 결과가 나오지 않았다.(기간별/장르별 변화가 다이나믹 하지 않았다)
2. 자연어처리의 고도화 문제로 인해 보다 다양하고 정확한 키워드 추출이 불가능했다.
3. 위 이슈들로 인해 다양한 전체적인(다양한) 분야를 아우르는 결과를 도출하기보다는 하나의 분야에 한해 분석을 하게 되었다.
4. 제목만으로 크롤링 했기 때문에 정확도 측면에서 미흡함.

호건:

NLP는 데이터사이언스 강의에서 알고자 하는 내용이 아니기에 대중적인 패키지 선까지만 이용하고 데이터 분석을 활용하는 방향으로 진행을 하였습니다.

~~한 과정을 통해 분석을 진행하였지만, 키워드가 반복되는 점 / 한정된 키워드가 나오는 점 등의 한계가 있어 유의미한 분석 결과가 나오지 않은 것 같아 아쉬움이 남습니다.

<향후 방향>

1. 추후 줄거리와 기사 내용도 크롤링을 해서 좀더 키워드 선정에 정확도를 높일 예정이다.
- 2.

NLP 전처리를 추가해볼까

<발표>

1. 안녕하세요. 데이터 사이언스 9조 프로젝트 발표를 맡은 김성윤입니다.
2. 본격적인 주제 소개에 앞서 가벼운 질문으로 시작하자면, 여러분들은 수업을 들으실 때, 종이 책과 pdf파일 중에 어떤 것을 사용하고 계신가요? 제가 느끼기로는 대부분의 사람들이 종이책보다 pdf파일을 사용하는 것 같습니다. 아는 선배가 이렇게 말합니다. '나 때는 두꺼운 전공책 들고 다니느라 죽는 줄 알았어~'
3. 이렇듯 저희는 디지털 시대로의 전환의 시기에 있습니다.
4. 이러한 전환의 시기에서 종이 책을 출판하는 출판회사들의 상황은 어떨까요? 종이 책을 덜 사용하니까 그만큼 심각하지 않을까요?
5. 예상과 다르게 최근 출판시장은 점점 성장하고 있습니다.
6. 세계 혁신 지수(GII)에 따르면 2022-2026년 세계출판시장 규모는 192억 달러의 증가가 전망된다고 합니다. 왜냐하면, 종이책이 줄어들지만, 그만큼 전자책의 전망이 좋기 때문입니다. 이는 총매출액을 통해서 확인할 수 있는데요. 대한출판문화협회의 '2022년 출판시장 통계' 보고서를 보면, 최근 출판회사들의 매출은 전년 대비 2.8%(1396억원) 정도로 소폭 증가했다는 것을 알 수 있습니다.
7. 하지만, 총매출액과 달리 출판시장의 영업이익은 전년 대비 38.7%(1312억원)이나 감소했습니다. 이게 무슨 말이나 함은 도서의 판매량은 많이 늘었지만, 실제로 출판사가 가져가는 돈은 줄어들었다는 것을 의미합니다. 마케팅과 인건비 등을 그 원인을 생각할 수 있죠. 또한, 총매출액과 관련 있는 판매량 역시 교육출판사만 높게

증가했다는 것입니다. **ruff**하게 생각하자면, 임시 책 때문에 총매출액이 많아보는 상황인 겁니다.

8. 지금까지 이야기를 정리하자면, 출판회사들은 계속 성장하는 출판시장 속에서 영업이익을 계속 늘려 나아가야 합니다.
9. 그러기 위해 총매출액이라도 높은 교육출판사와 달리 경제와 과학과 같은 지식을 다루는 출판사는 위험한 상황
10. 그러면 어떻게 영업이익을 증가할 것이냐!?
11. 그래서 준비했습니다. ‘독서 트렌드와 영향 요인 파악’, 저희의 목표입니다.
12. 저희 기대효과는 이렇습니다. “독서 트렌드와 영향 요인을 파악하여 자동화 한다면, 이를 마케팅 전략에 활용하여 영업손실에 해당하는 마케팅 비용과 인건비 등을 최소화할 수 있을 것이다.” 저희는 이러한 ‘독서 트렌드’와 ‘영향 요인’을 파악하기 위해서 다음 세 가지의 데이터를 분석해보았습니다.
13. ‘도서관 데이터’, ‘서점 데이터’, ‘뉴스 기사 수 데이터’. 독서 트렌드의 영향 요인에는 다양한 요인이 존재합니다. 그 중 ‘(뉴스 기사로 인해) 특정 소재가 이슈화된다면, 이는 독서 트렌드에 영향을 미칠 것이다.’가 저희의 가설이었습니다.
14. 각각의 데이터는 도서관 정보나루, 알라딘, 구글 뉴스 크롤링을 통해서 얻을 수 있었습니다.
15. 본격적으로 분석 과정을 살펴보면, 크게 3단계로 생각할 수 있습니다. **step1** 도서관 데이터와 서점 데이터 중 베스트 셀러와 대출 순위가 높은 상위 1000개의 데이터 중에서 ‘경제’에 해당하는 책의 제목을 받아옵니다. **step2** 이 제목들 중 **tf-idf**를 활용해서 키워드를 뽑아냅니다. **tf-idf**를 활용하면 단어의 상대적인 중요도를 파악할 수 있어서 점수가 높은 순서대로 키워드라고 해석했습니다. **step3** 이 키워드를 구글 검색창에 검색하고 뉴스 기사의 개수가 몇 개인지를 파악했습니다. 이 과정을 2020-2022년 3년 간 각 월별로 진행했습니다.
16. 그리고 다음은 저희가 얻은 시각화 결과에서 유의미하게 데이터를 선정해보았습니다.
17. <투자> - 대중적인 키워드이기 때문에 항상 높은 퍼센트를 보였다. 대체적으로 **tf-idf** 점수와 기사 수 퍼센트의 양상이 같은 모습을 보였다. <부동산> - 임대차법 발의 및 의결 시기인 20년 7월 이전으로 높은 기사 수와 점수를 보였다. 해당 법안의 효력이 발휘되는 시점인 21년도 하반기 말부터 다시 기사 수와 점수가 오르는 모습을 보였다. <주식> - 투자와 같이 대중적인 키워드이기에 거의 높은 수치를 유지하고 있다. 상대적으로 경제 관심도가 떨어지는 연말 연초에는 기사와 점수 모두 하락하는 모습을 보였다. <부자> - 부자라는 키워드는 너무 추상적인 키워드라 뵈까 했지만 경제를 상징하는 단어라고 생각해 같이 분석을 해보았다. 부자 키워드의 특징은 기사 수 퍼센트에 해당하는 점수 시기가 약 두달가량 뒤로 밀린 형태가 나왔다. 키워드가 추상적이라 발생하는 특징인듯 하다. <투자> - 부동산, 주식, 가상화폐 등 경제의 대다수를 차지하는 키워드가 전부 투자 키워드와 관련이 있기에 항상 높은 관심도와 점수를 보였다. 마찬가지로 추상적인 키워드이기에 퍼센트와 점수 간극이 한달 정도 발생하는 모습이다.
18. 그래서 결론을 내리자면 ~ 라고 할 수 있습니다.
19. 이를 통해서 저희의 가설이 ~다는 것을 알 수 있었습니다.
20. 저희에게는 4가지 한계점이 존재합니다. ~ 입니다.
21. 이를 해결하기 위해 다음과 같이 진행할 예정입니다.~ 입니다.
22. 여기까지 저희의 핵심 발표를 마치고요.
23. 부록으로 팀원별 역할입니다.
24. Q&A

=> 읽고 말했을 때, 6분 30초 이내

tf-idf 점수 변동, 기사 수 변동을 이용하면 특이점 파악에 도움될거예요

<결과 분석>

각 도서관/서점 사회과학 분류중 경제 분야 서적의 제목을 **Konlpy** 형태소 분석 모델?
패키지?를 통해 키워드를 추출했습니다. 그 결과 **20~22** 대부분의 기간 동안
'주식','투자','부동산'등의 공통적인 키워드를 발견할 수 있었고, 저희는 이 키워드의 tf-idf
점수 변화와 기사 수 등을 활용하여 특이점을