

프로젝트 명 : 라이브 맵 (library map / live map)

도서관 특징 분석

발 표 자 이 혜 원

AI 전문인력 양성과정
청년취업사관학교 알파코

CONTENTS

01

프로젝트 소개

- 주제 선정 이유
- 사용한 데이터
- 개발 환경

02

역할 소개

- 역할 소개

03

데이터 전처리

- 데이터 확인
- 데이터 전처리
- 데이터 이상치

04

도서관 EDA

- 기본 통계
- 데이터 시각화

05

마무리

- 분석 결과 / 요약
- 참고 사이트

01

프로젝트 소개

- 주제 선정 이유
- 사용한 데이터
- 개발 환경

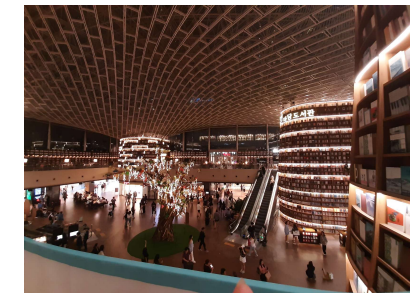
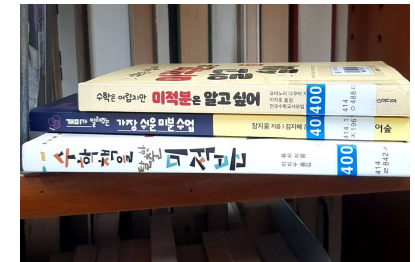
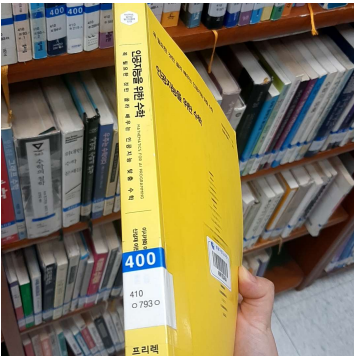
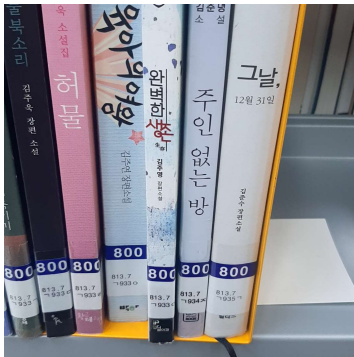
01

“ 주제 선정 이유 ”

도서관이 많은 곳 주변에 살고 싶다!!
하지만 도서관이 어디에 분포되어 있는지 잘
시각화 된 자료가 없는 것 같다.

도서관에 관한 궁금증도 풀고 싶다.
어떤 연령대가 도서관을 많이 갈까
어떤 책을 많이 빌릴까?
언제 도서관에 사람이 비교적 적을까?

도서관에 관한 궁금증을 통계수치와 시각화를
통해 확인해보자



01

“ 사용한 데이터 ”

전국도서관표준데이터: 공공데이터 포털 데이터셋 활용
<https://www.data.go.kr/data/15013109/standard.do>

연령별 스테디 대출 분석, 주제별/세대별 다대출 도서 :
도서관 정보나루 > 데이터 신청 > 이용자별 테마 활용
<https://www.data4library.kr/userThemaDataL>

DATA 공공데이터포털
GO . KR

데이터찾기

국가데이터맵

데이터요청

데이터활용

정보공유

이용안내

정보공유>

개발자 네트워크

데이터 공유하기

공지사항

자료실

문의하기

포털뉴스

표준데이터 상세



>URL 복사

CSV 전국도서관표준데이터

공공데이터 제공 표준데이터 속성정보(허용값, 표현형식/단위 등)는 [공공데이터 제공 표준] 전문을 참고하시기 바랍니다.
(공공데이터포털>정보공유>자료실) 각 기관에서 등록한 표준데이터를 취합하여 제공하기 때문에 갱신주기는 개별 파일마다
다릅니다.(기관에서 등록한 데이터를 취합한 것으로 개별 파일별 갱신시점이 다름)



5



0

□ 관심

표준데이터 기본정보

메타데이터 다운로드

오류신고 및 문의

제공범위(대상)	「도서관법」 등에 따라 지방자치단체에서 관리하고 있는 도서관에 관한 정보(단, 사립으로 운영 중인 도서관은 해당 지방자치단체에 신고를 한 경우에만 포함 / 병영도서관, 교도소도서관은 제공범위(대상)에서 제외)		
관련법령	도서관법, 작은도서관 진흥법, 학교도서관진흥법, 독서문화진흥법, 지방자치단체 도서관 및 독서문화 진흥 조례 등		
소관기관	문화체육관광부	제공기관	지방자치단체, 교육청
표준데이터셋제공시스템	해당없음	다운로드 수	39305
갱신주기	연간	수정일	2023-08-10
분류체계	교육 - 교육일반	키워드	도서관,마을문고,서적

데이터 상세정보



분류체계	교육 - 교육일반	제공기관	인천광역시교육청 인천광역시교육청계양도서관
관리부서명	정보자료과	관리부서 전화번호	032-540-4434
업데이트 주기	연간	차기 등록 예정일	2024-07-05
매체유형	텍스트	전체 행 수	
확장자	CSV	다운로드(바로가기) 횟수	37
등록일	2020-07-07	수정일	2023-07-11
데이터 다운로드	<div>CSV</div>		
비용부과유무	무료	비용부과기준 및 단위	
이용허락범위	이용허락범위 제한 없음		

닫기

01

DATA 공공데이터포털

데이터 소개정보

요청 데이터	요청일시	완료일시	진행상태
연령별 스테디 대출 분석 2013-01-01 ~ 2023-08-15	2023-08-15 10:36	2023-08-15 10:42	 다운로드
연령별 스테디 대출 분석 2021-08-01 ~ 2023-08-15	2023-08-15 10:22	2023-08-15 10:31	 다운로드
2021-08-01 ~ 2023-08-15	2023-08-15 02:01		오류
주제별/세대별 다대출 도서 2021-08-01 ~ 2023-08-15	2023-08-15 02:00	2023-08-15 02:21	 다운로드

01

“ 개발 환경 ”



사용한 파이썬 라이브러리



re(정규표현식)



WordCloud

PIL

02 역할 소개

- 역할 소개

02

“ 역할 소개 ”



이혜원

담당 : 기획, 데이터 수집, 데이터 전처리, 데이터 시각화, EDA,
PPT 작성, 발표

03 데이터 전처리

- 데이터 확인
- 데이터 전처리
- 데이터 이상치

03

“ 데이터 확인 ”

1) 전국도서관표준데이터

	A	B	C	D	E	F	G	H	I	J
1	도서관명	시도명	시군구명	도서관유형	휴관일	평일운영시작시각	평일운영종료시각	토요일운영시작시각	토요일운영종료시각	공휴일운영시작시각
2	이석영뉴	경기도	남양주시	공공도서관	둘째·넷째	9:00	22:00	9:00	18:00	9:00
3	원덕도서관	강원도	삼척시	공공도서관	토요일+일	9:00	18:00	0:00	0:00	0:00
4	가평군 한	경기도	가평군	공공도서관	1월 1일+설	7:00	23:59	7:00	23:59	7:00

K	L	M	N	O	P	Q	R	S	T
공휴일운영종료시각	열람좌석수	자료수(도서)	자료수(연속간행물)	자료수(비도서)	대출가능권수	대출가능일수	소재지도로명주소	운영기관명	도서관전화번호
18:00	615	48913	111	643	20	14	경기도 남양주시 호사단법인 문	031-595-8390	
0:00	4	78766	50	160	10	30	강원도 삼척시 원	033-572-4819	
23:59	119	140341	44	5536	7	14	경기도 가평군 가	031-580-4041	

U	V	W	X	Y	Z	AA	AB	AC	AD
부지면적	건물면적	홈페이지주소	위도	경도	데이터기준일자	제공기관코드	제공기관명		
1843	4877	http://lib.nyj.g	37.65816	127.2995	2023-04-30	3990000	경기도 남양주시		
766	364		37.17685	129.3355	2023-06-20	4241000	강원특별자치도 삼척시		
7485	3602	www.gaplib.g	37.82969	127.5064	2023-06-21	4160000	경기도 가평군		

03

“ 데이터 전처리 ”

1) 전국도서관표준데이터 - 결측치 처리

위도, 경도가 없다.

도서관명	소재지도로명주소	위도	경도
사천시어린이도서관	경상남도 사천시 사남면 조동길 49-30		
대림동산작은도서관	경기도 안성시 공도읍 대림동산길 38		
양성작은도서관	경기도 안성시 양성면 읍내길 21		
미양작은도서관	경기도 안성시 미양면 양지길 29		
안성시 진사2길 도서관	경기도 안성시 공도읍 진사2길 70		
안성시 안성면 안성초등학교	경기도 안성시 안성면 안성초 422		

03

“ 데이터 전처리 ”

도서관명	
사천시어린이	경기
대림동산직	경기
양성작은도	경기
미양작은도	경기
안성시 진	경기

1 #NaN값 확인 - 위도, 경도가 없는 경우가 많음 -> 소재지 도로명 주소는 있음	
2 data.isnull().sum()	
도서관명	0
시도명	0
시군구명	0
도서관유형	0
휴관일	0
평일운영시작시간	0
평일운영종료시간	0
토요일운영시작시간	0
토요일운영종료시간	0
공휴일운영시작시간	0
공휴일운영종료시간	0
열람좌석수	0
자료수(도서)	0
자료수(연속간행물)	0
자료수(비도서)	0
대출가능권수	0
대출가능일수	0
소재지도로명주소	0
운영기관명	668
도서관전화번호	233
부지면적	1610
건물면적	481
홈페이지주소	1422
위도	258
경도	247
데이터기준일자	0
제공기관코드	0
제공기관명	0
dtype: int64	

경도

03

```
1 print(len(nan_df))      # 위경도가 없는 데이터 개수  
2 print(len(result_list))  
3 print(result_list.count('')) # 위경도를 찾지 못한 데이터
```

```
258  
258  
46
```

```
데이터기준일자      0  
제공기관코드        0  
제공기관명          0  
dtype: int64
```


03

“ 데이터 확인 ”

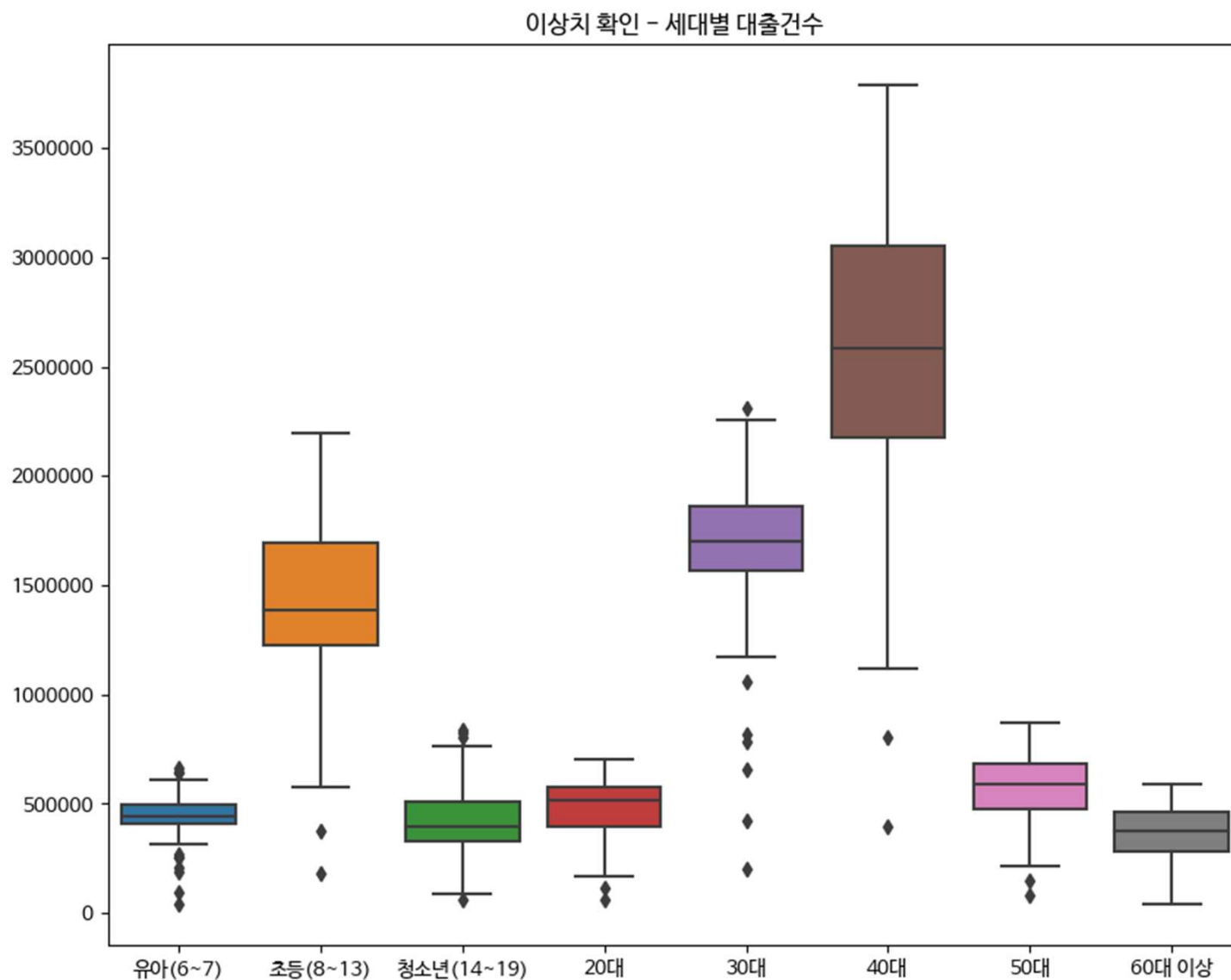
2) 연령별 스테디 대출 분석
2013-01월 ~ 2023-07월 데이터 사용

	A	B	C	D	E	F	G	H	I	J
1	월	영유아 (0~5)	유아(6~7)	초등 (8~13)	청소년 (14~19)	20대	30대	40대	50대	60대 이상
2	2013-01	305,064	426,040	1,819,538	806,154	703,926	1,678,550	2,204,407	451,948	237,981
3	2013-02	255,645	335,425	1,225,038	530,587	541,512	1,344,531	1,690,606	387,854	204,398
4	2013-03	327,166	428,211	1,349,098	430,425	552,441	1,650,861	1,965,168	424,935	222,853
5	2013-04	325,880	415,621	1,138,084	311,233	532,388	1,571,194	1,738,171	390,536	208,388
6	2013-05	289,157	358,086	994,972	360,298	496,881	1,402,621	1,589,619	363,065	193,884
7	2013-06	313,489	388,502	1,033,750	338,562	515,924	1,448,666	1,608,310	355,912	199,111
8	2013-07	353,294	440,556	1,313,266	559,618	601,386	1,615,899	1,888,900	413,884	221,580
9	2013-08	391,022	510,644	1,599,126	566,786	596,781	1,773,948	2,056,664	418,036	228,789
10	2013-09	369,525	423,758	1,103,808	317,810	518,759	1,544,894	1,695,568	368,408	200,828

03

데이터 이상치

2) 연령별 스테디 대출 분석



03

“ 데이터 확인 ”











3) 주제별/세대별 다대출 도서
2021-08월 ~ 2023-07월 데이터 사용

	A	B	C	D	E	F	G	H	I	J
1	월	영유아 (0~5)	유아(6~7)	초등 (8~13)	청소년 (14~19)	20대	30대	40대	50대	60대 이상
2	2021-08	11,292	18,043	76,257	33,311	29,086	112,569	188,339	54,675	28,423
3	2021-09	10,397	16,448	62,769	18,969	28,011	105,861	167,365	51,204	27,429
4	2021-10	10,473	16,186	59,797	19,842	28,498	104,228	163,978	51,301	26,990
5	2021-11	10,976	16,028	55,854	19,193	27,463	104,987	157,734	48,905	26,645
6	2021-12	10,686	16,479	54,194	20,582	25,238	96,765	150,707	48,191	27,597
7	2022-01	7,466	12,333	75,090	28,144	29,348	91,360	186,460	59,251	32,638
8	2022-02	7,108	11,720	65,913	20,995	25,329	85,547	164,955	52,835	30,583
9	2022-03	8,287	13,480	65,890	21,220	27,534	97,262	181,299	59,081	32,563
10	2022-04	9,025	14,041	65,264	17,685	26,696	97,475	178,666	57,830	31,130

03

“ 데이터 확인 ”

3) 주제별/세대별 다대출 도서
2021-08월 ~ 2023-07월 데이터 사용

	A	B	C	이름	H	I	J
1	월	영유아 (0~5)	유아(6		40대	50대	60대 이상
2	2021-08	11,292	18,	 이용자테마데이터_기술과학-세대별대출...	188,339	54,675	28,423
3	2021-09	10,397	16,	 이용자테마데이터_문학-세대별대출분석	167,365	51,204	27,429
4	2021-10	10,473	16,	 이용자테마데이터_사회과학-세대별대출...	163,978	51,301	26,990
5	2021-11	10,976	16,	 이용자테마데이터_언어-세대별대출분석	157,734	48,905	26,645
6	2021-12	10,686	16,	 이용자테마데이터_역사-세대별대출분석	150,707	48,191	27,597
7	2022-01	7,466	12,	 이용자테마데이터_예술-세대별대출분석	186,460	59,251	32,638
8	2022-02	7,108	11,	 이용자테마데이터_자연과학-세대별대출...	164,955	52,835	30,583
9	2022-03	8,287	13,	 이용자테마데이터_종교-세대별대출분석	181,299	59,081	32,563
10	2022-04	9,025	14,	 이용자테마데이터_철학-세대별대출분석	178,666	57,830	31,130
				 이용자테마데이터_총류-세대별대출분석			

03

“ 데이터 확인 ”

3) 주제별/세대별 다대출 도서
2021-08월 ~ 2023-07월 데이터 사용

영유아(0~5)

유아(6~7)

초등(8~13)

청소년(14~19)

20대

30대

40대

50대

60대 이상

03

“ 데이터 확인 ”

3) 주제별/세대별 다대출 도서
2021-08월 ~ 2023-07월 데이터 사용

	A	B	C	D	E	F
1	순위	서명	저자	출판사	출판년도	권
2	1	4~7세보다 중요한 시기는 없습니다 :아이의 정서와 인지 발달을 키우는 결정적 시기	이임숙 지음	카시오페아	2021	
3	2	어떻게 말해줘야 할까 :오은영의 현실밀착 육아회화	글: 오은영 ;그림: 차상미	김영사	2020	
4	3	소중해 소중한 나도 너도 :3-7세 영유아와 어른들을 위한 첫 성 교육 그림책	엔미 사키코 글 ;권남희 옮김	주니어RHK	2022	
5	4	공사하는 자동차	고요세 준지	터치아트	2018	
6	5	(우리가 꼭 알아야 할) 세상 모든 중장비	탈것발전소 지음	주니어골든벨	2019	
7	6	(안녕과 질서를 지켜주는) 세상 모든 경찰차	안명철 글	주니어골든벨	2019	8

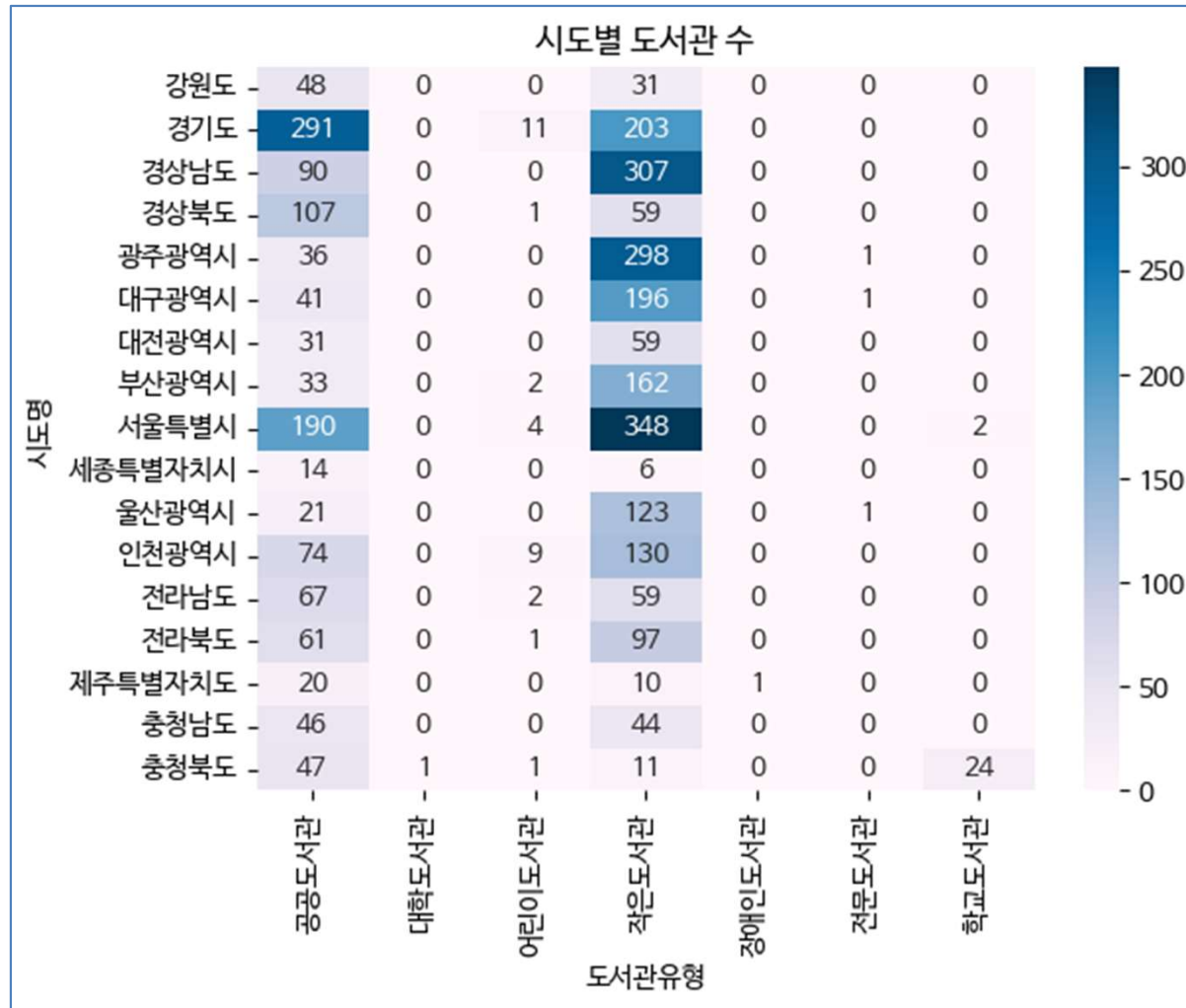
G	H	I	J
ISBN	ISBN부가기호	KDC	대출건수
9791190776820	03590	598.122	765
9788934986652	13590	598.1	655
9788925578774	77830	598.5	541
9791187936152			486
9791158064006	77550	550	470
9791158063900	77550	550	468

04 도서관 EDA

- 기본 통계
- 데이터 시각화

도서관 EDA

1) 시도별 도서관 수



시도별로 도서관 수가 차이가 난다.

공공도서관은 경기도와 경상북도와

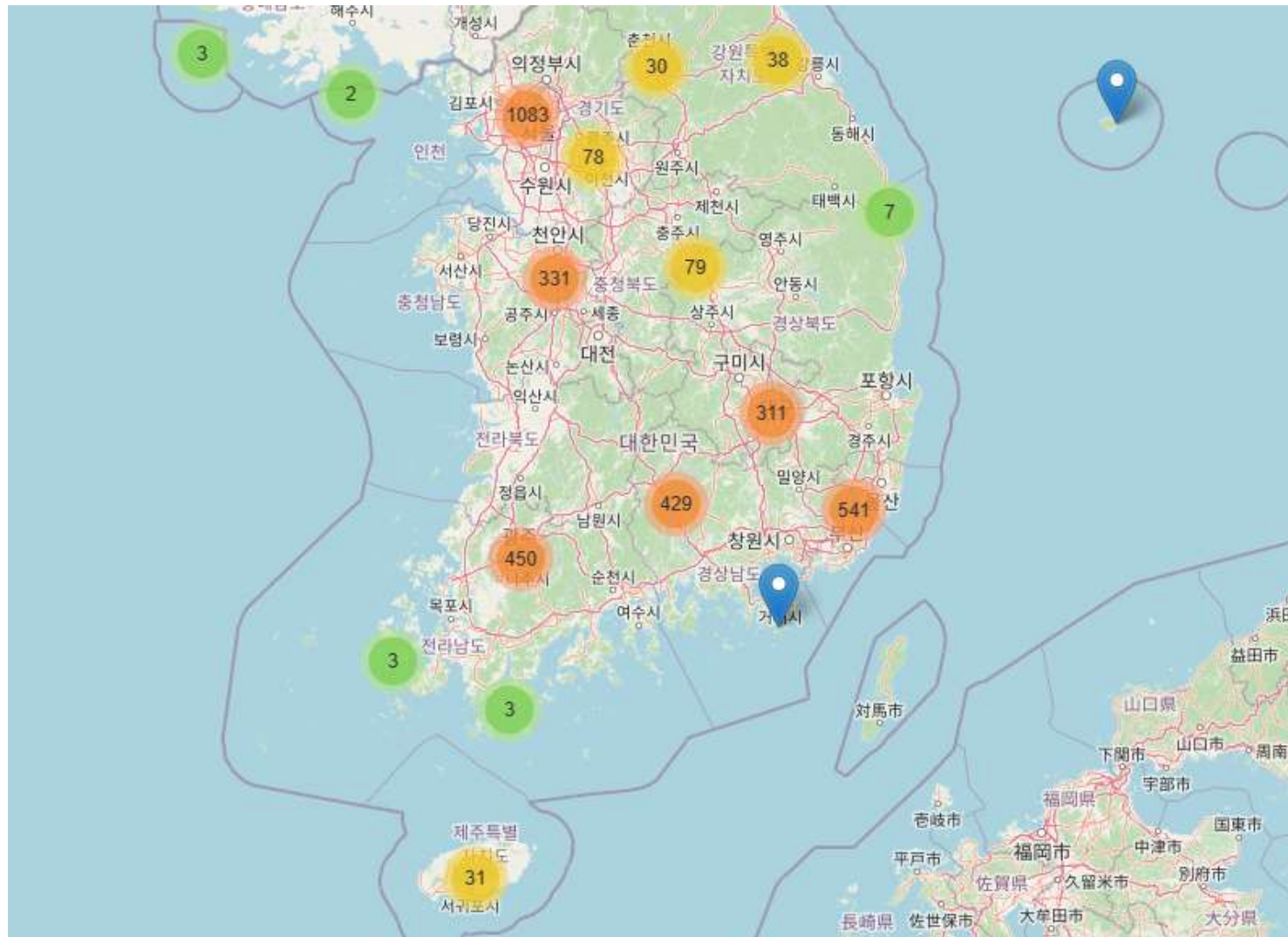
서울특별시에 많이 있다.

공공도서관의 경우 서울보다 경기도가 더 많은 것을 확인하였다.

04

도서관 EDA

1) 시도별 도서관 수



시각화 결과 배포 페이지: live-map.netlify.app

04

도서관 EDA

2) 연령대별 평균 대출건수

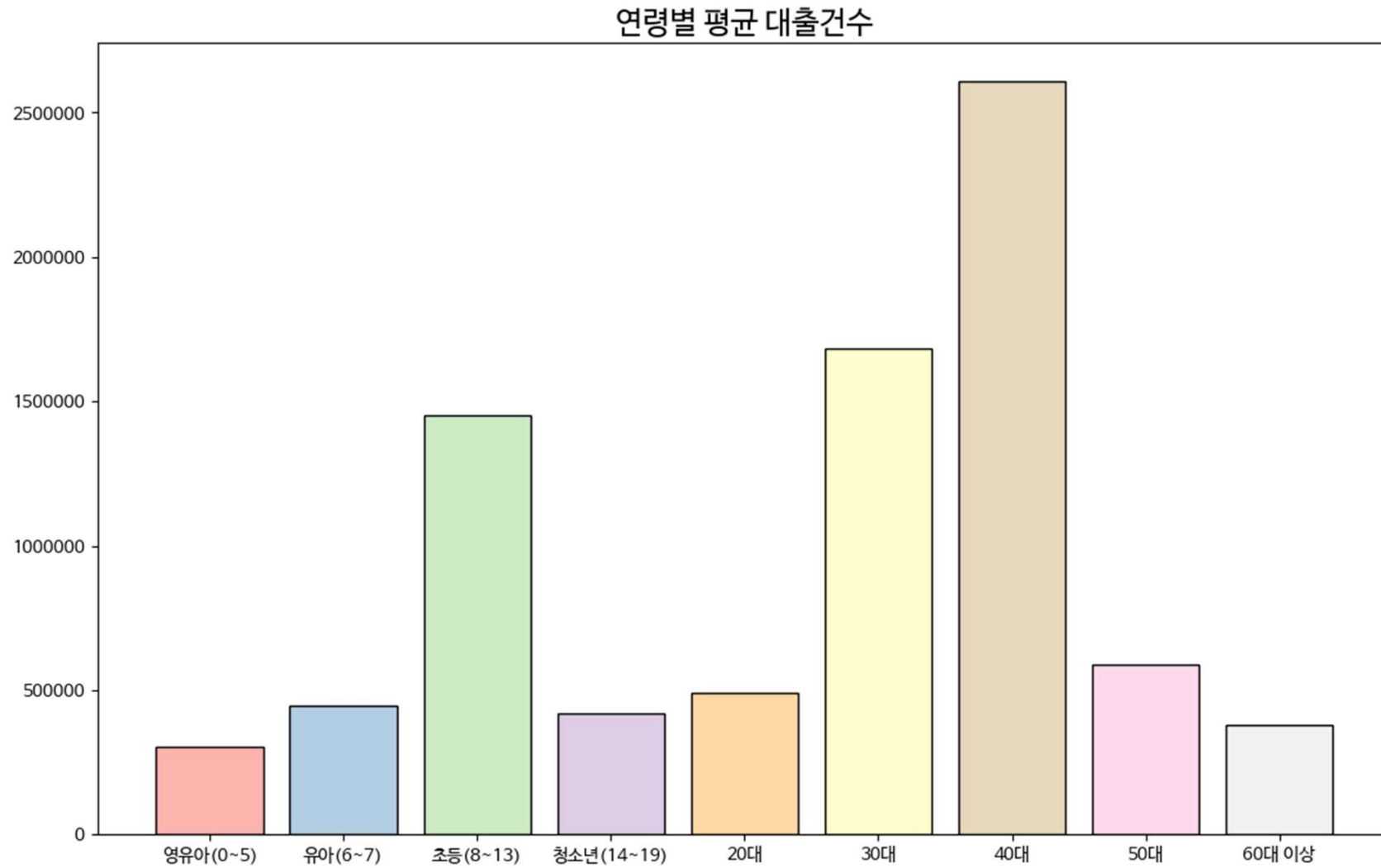
기본 통계를 내봤더니 이렇게 나왔습니다.

이대로 보기엔 한눈에 들어오지 않기 때문에 시각화를 해봤습니다.

	영유아(0~5)	유아(6~7)	초등(8~13)	청소년(14~19)	20대	30대	40대	50대	60대 이상
count	127.00000	127.00000	127.00000	127.00000	127.00000	127.00000	127.00000	127.00000	127.00000
mean	305808.82677	444018.40157	1442924.30709	423590.60630	492600.50394	1690361.40945	2579929.46457	579094.06299	371826.45669
std	68110.99996	93488.47218	334636.13114	149244.03879	116682.67709	309401.66236	630154.29018	145230.24055	111631.70911
min	23911.00000	39606.00000	184596.00000	58496.00000	61230.00000	201352.00000	395885.00000	81178.00000	41608.00000
25%	269310.00000	406384.00000	1228976.50000	330158.50000	403639.00000	1570573.00000	2184138.50000	480155.50000	285692.50000
50%	313696.00000	441345.00000	1394626.00000	395258.00000	514945.00000	1700550.00000	2587969.00000	589290.00000	379557.00000
75%	352720.00000	494287.00000	1696098.00000	509646.50000	579569.50000	1861523.00000	3054008.00000	684562.50000	461098.50000
max	451535.00000	660816.00000	2197574.00000	840170.00000	704374.00000	2309395.00000	3787137.00000	868545.00000	592792.00000

도서관 EDA

2) 연령대별 평균 대출건수



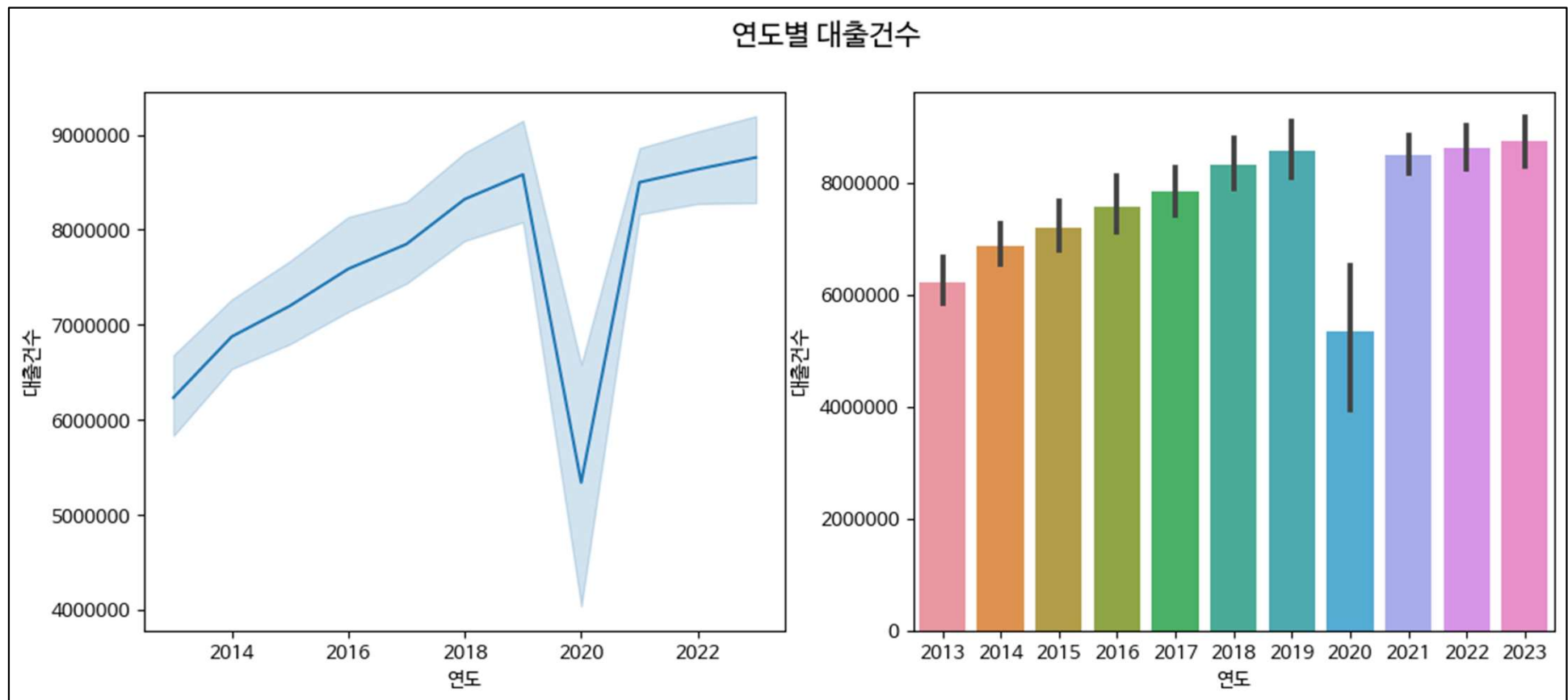
04

도서관 EDA

3) 연도별 대출건수

연도별 대출건수를 확인해봤더니 2020년에 코로나로 인해 줄어든 것이라고 예측이 되고 계속 오르
고 있는 것을 확인할 수 있습니다.

하지만 대출 데이터를 공개한 도서관이 늘어서 대출 건수도 늘었을 가능성도 있습니다.



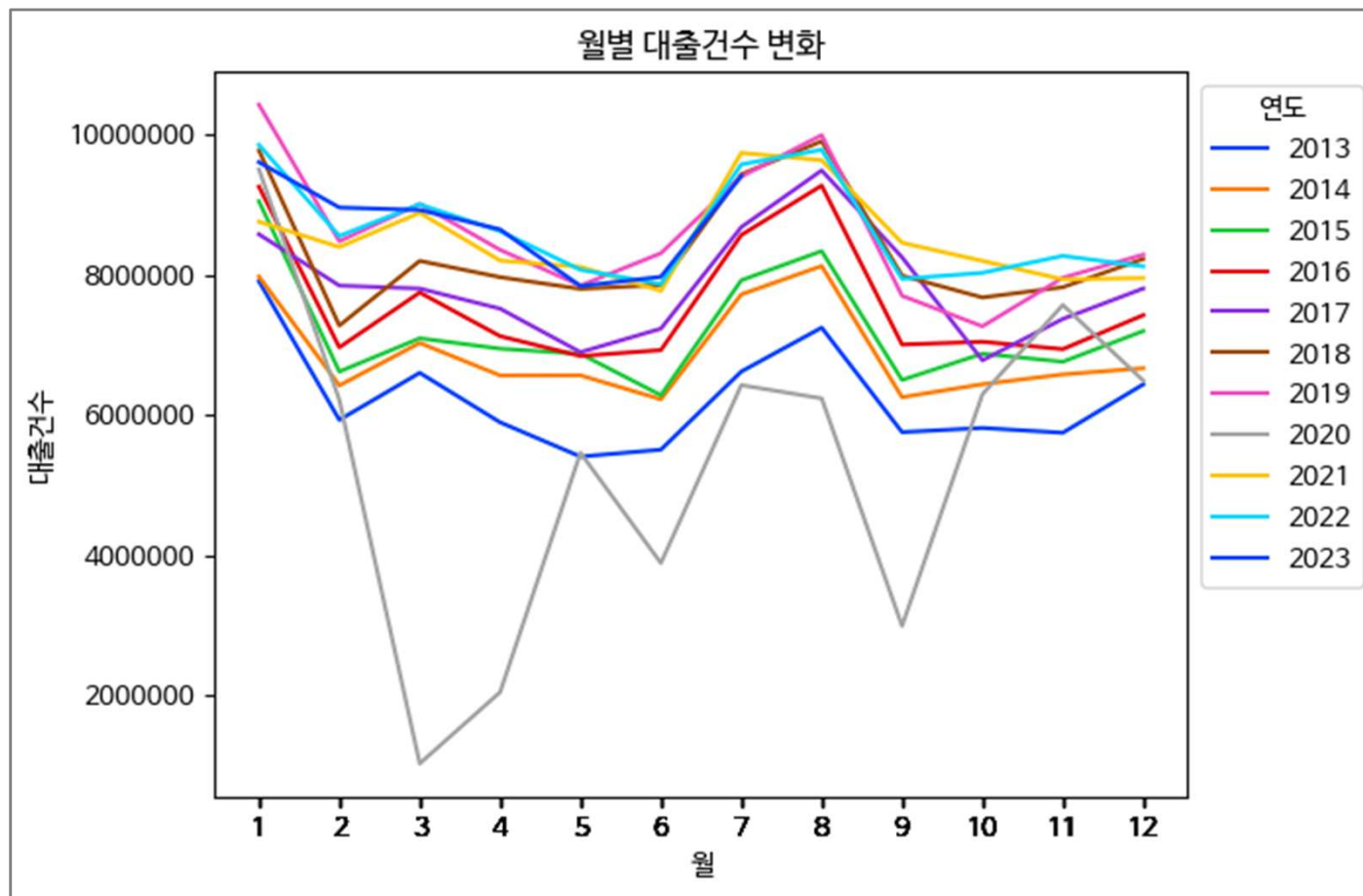
도서관 EDA

4) 월별 대출건수 변화 선 그래프

특수상황인 코로나 2020년을 제외하고는 대출건수가 모두 같은 패턴인 것을 발견하였습니다.

주로 도서관 이용자들은 1월, 7월, 8월에 많이 대출을 하는 것을 알 수 있고

해당 월에 이용자들이 도서관에 많이 방문한다고 추측해볼 수 있습니다.

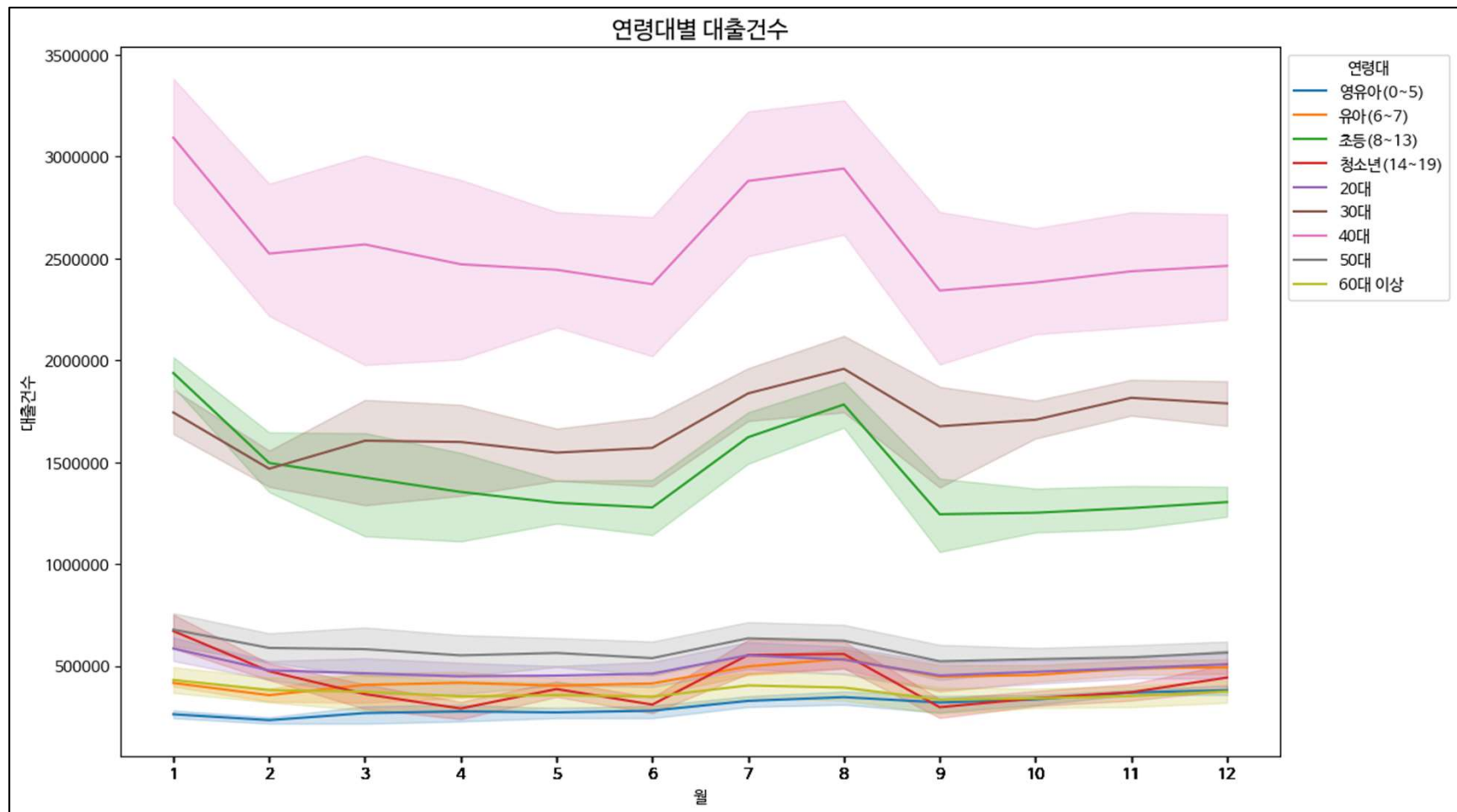


도서관 EDA

5) 월별 연령대별 평균 대출건수

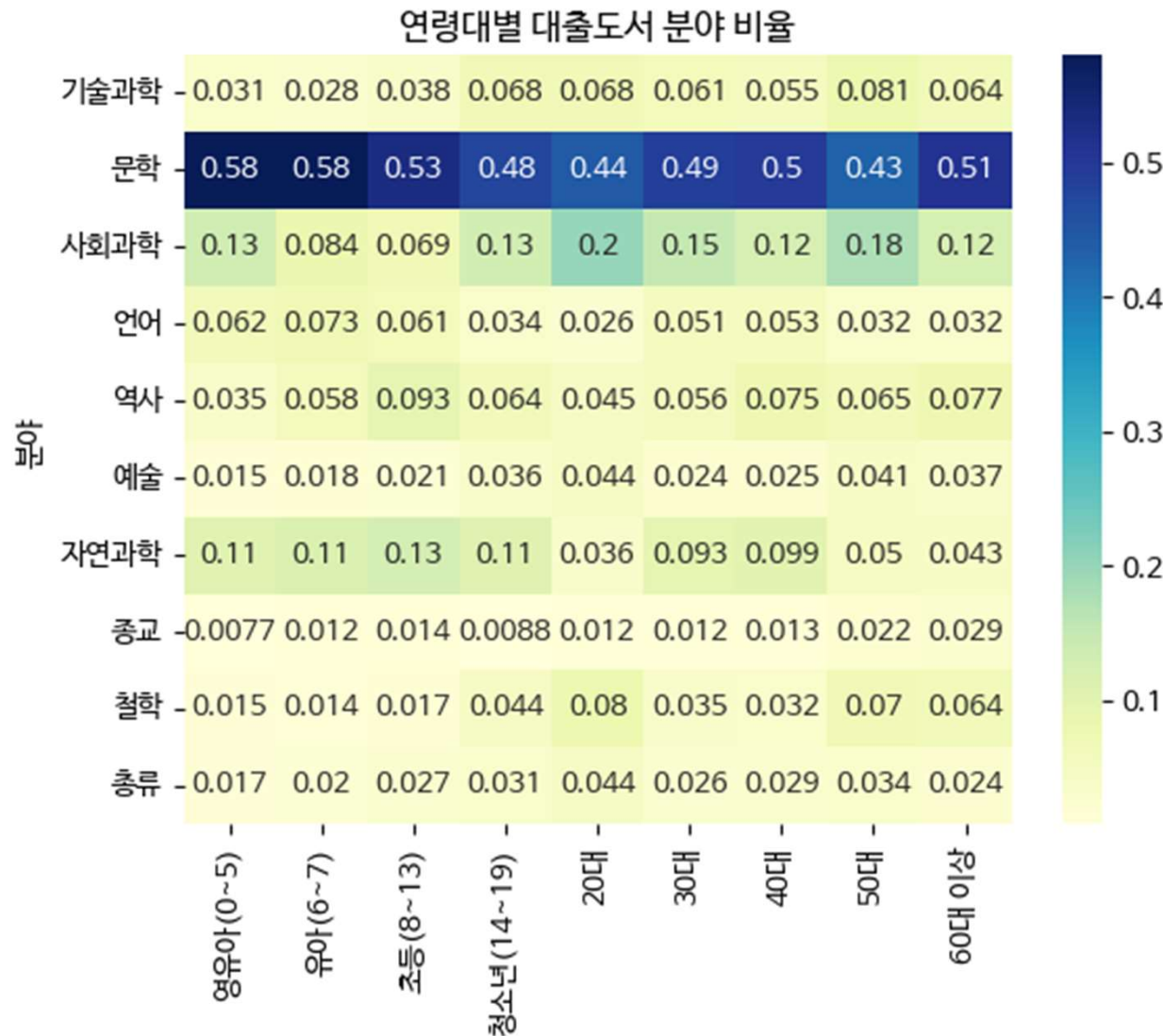
한 연령대가 유독 많이 방문하는 달이나 적게 방문하는 달은 딱히 없고, 앞서 확인한 그래프처럼

1,7,8월에 대출을 많이 하고 비슷하게 분포되어 있습니다.



도서관 EDA

6) 연령대별 대출도서 분야 비율 2021-08월 ~ 2023-07월 데이터



연령대별로 비율을 구해서 히트맵으로 시각화하였더니 세대 불문하고 문학을 가장 많이 대출하는 것을 알 수 있습니다.

도서관 EDA

7) 많이 빌린 도서 – 워드클라우드
2021-08월 ~ 2023-07월 데이터

세대별, 주제별 많이 대출한 도서 각각 상위 5권 썩의 데이터를 사용하였고 대출건수를 빈도수로 하여 워드클라우드를 그렸습니다. 시리즈물 책이 책 이름이 같을 경우 빈도수를 합산하여 그렸습니다. 설민석 한국사 책, 그리스 로마 신화, 마법천자문, 불편한 편의점, 흔한남매, 김 부장 이야기 책이 2년간 많이 대출된 것을 알 수 있습니다.

05 마무리

- 분석 결과/요약
- 참고 사이트

“ 분석 결과/요약 ”

도서관 수가 시도별로 차이가 나고 공공도서관은 **경기도와 경상북도와 서울특별시에** 많이 있다.

40대, 30대, 초등학생이 많이 빌리고 **20대**는 그에 비해 적게 대출을 한다는 것을 알게 되었다.

1월, 7월, 8월에 많이 대출을 한다.

전연령 사람들이 여러 분야 중에 **문학** 주제의 책을 제일 많이 대출한다.

이러한 분석 결과를 도서관 홍보전략에 사용하여도 좋을 것 같다고 생각이 들었습니다.

“ 참고 사이트 ”

1. 정규표현식 예제 - <https://nachwon.github.io/regular-expressions/>
2. 정규표현식 테스트 - <https://regex101.com/r/uufrWr/1>
3. 주소 경위도 확인 - <https://address.dawul.co.kr/index.php>
4. 주소 경위도 확인 - <https://www.google.co.kr/maps>
5. matplotlib 한글깨짐 이슈 -
<https://www.whatwant.com/entry/matplotlib-hangul-colab-local>
6. folium 사용 - <https://teddylee777.github.io/visualization/folium/>
7. matplotlib 색 -
<https://matplotlib.org/stable/tutorials/colors/colormaps.html>
8. seaborn 공식 사이트 (색상표) -
https://seaborn.pydata.org/tutorial/color_palettes.html
9. 피피티 템플릿 다운 - <https://minheeblog.tistory.com/144>

THANK
YOU

발 표 자 이 혜 원