

# 서울시 성북구 공공도서관 별 ‘인기 아동 도서’와 유사도 높은 ‘미대출 아동 도서’ 추천 시스템

(도서관 정보나루 제공 데이터 활용)

6조  
김사무엘  
김외솔  
서승우

# 프로젝트 개요

1  
문제상황

2  
대안

# 1 문제상황

- 도서관 별로 존재하는 ‘**미대출 도서**’ (도서관 개관 이후 한 번도 대출이 없었던 도서)
- 미대출 도서 대책은 ‘도서관 정보나루’ 빅데이터 사업의 주요 주제



## ▶ 활용 데이터

### 미대출 장서 분석

## ▶ 분석 내용

도서관에서 한 번도 대출이 되지 않았던 미대출 장서수와 미대출 장서 비율을 주제분야별로 파악하고, 세부주제별 미대출 장서 분포 분석

- 미대출 장서 분석을 통한 대상별, 주제별 미대출 도서 분석
- 기간(최근 5년, 3년, 1년)별 미대출 장서 분석

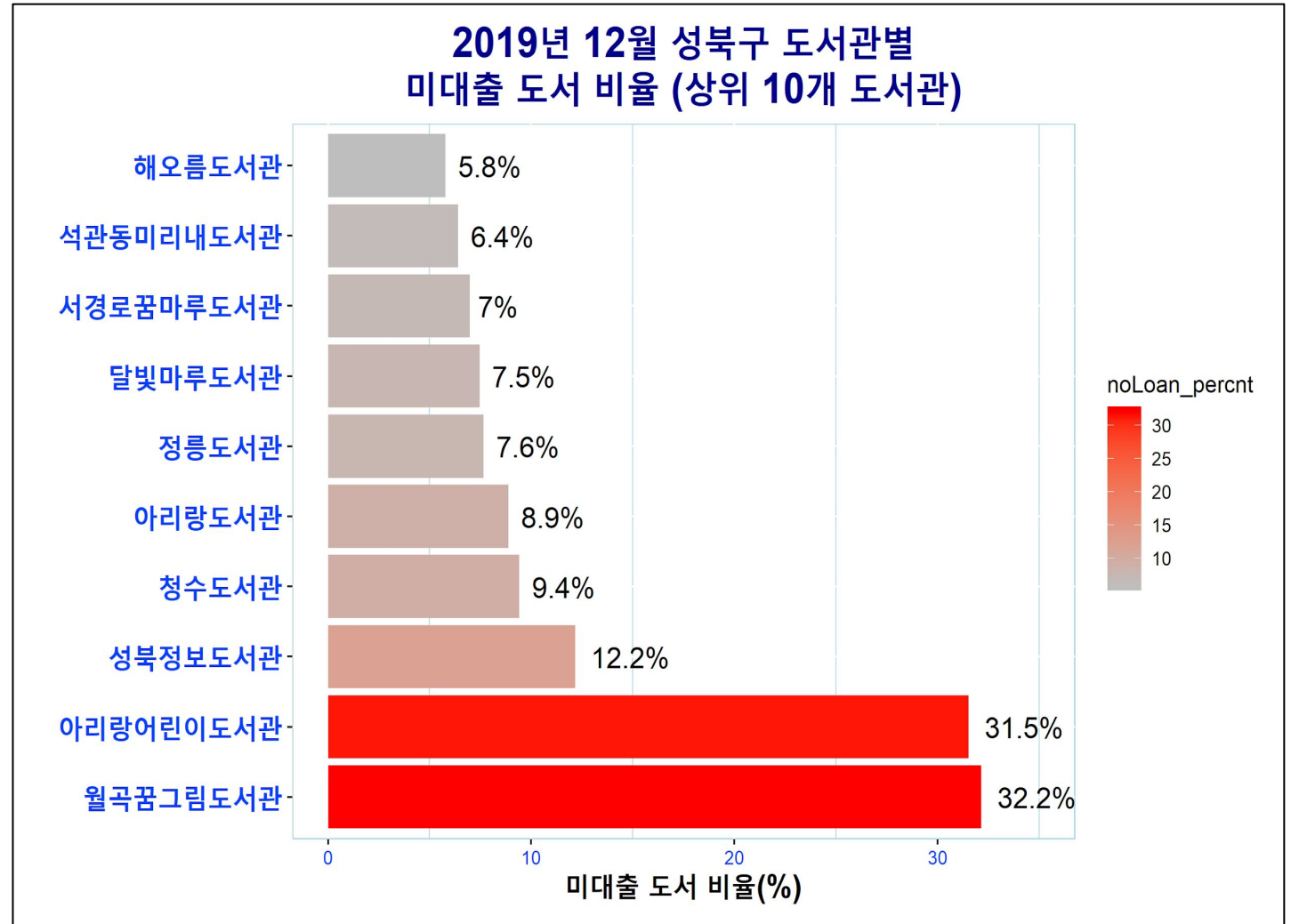
## ▶ 주요 사례

미대출 장서 활용을 위한 독서 이벤트 및 프로그램 개발

- 미대출장서활용을위한'책속의 보물 찾기' 행사 진행
- 미대출 장서 중심의 테마 서가 마련

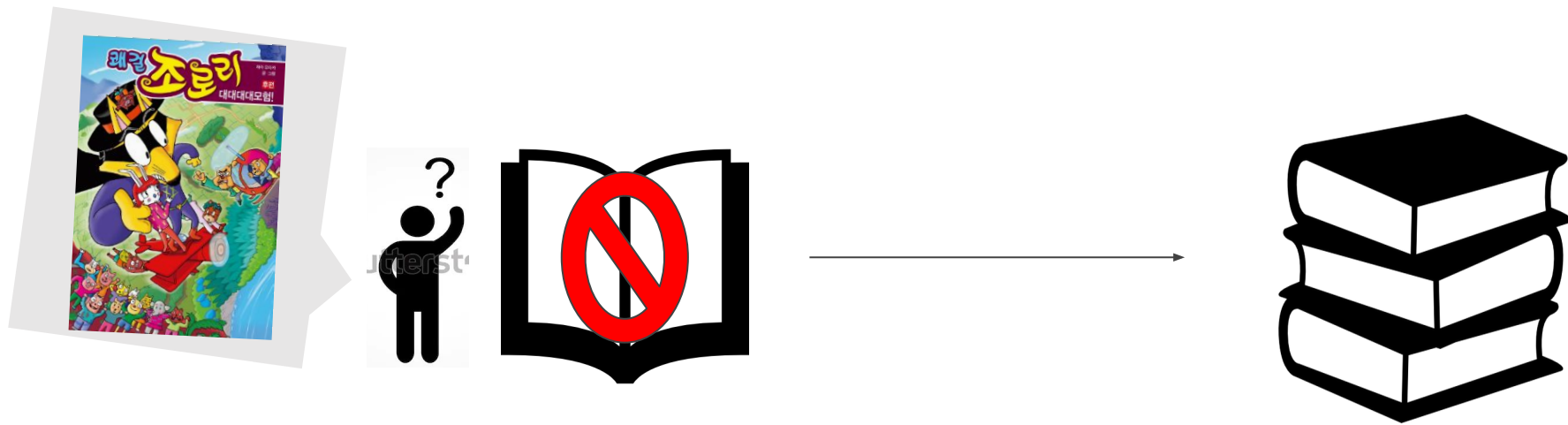
# 1 문제상황

그럼에도 불구하고,  
여전히 발견되는  
미대출 도서들



## 2 대안 : 도서 추천 시스템 → 미대출 도서 활성화

이용자가 찾는 인기도서가 **대출 중인 경우**,  
미대출 도서 중 **컨텐츠 유사도가 높은 도서**를 추천 !



※ 분석 범위는 ‘월곡꿈그림 / 아리랑어린이 도서관’, 인기 도서 목록은 ‘상위 10권’으로 한정

## 분석 내용 소개

1

데이터 수집,  
전처리

2

키워드 기반의  
유사도 분석

3

추천 도서  
목록 선정

사용 데이터

# 1 데이터 수집 및 전처리

서울시 성북구 11개 도서관의

2019년 12월 ‘도서관별 장서 / 대출 데이터(csv)’ 크롤링

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	관리번호	도서명	저자	출판사	발행년도	ISBN	세트 IS	부가기호	권	주제분류번호	도서권수	대출건수	등록일자
2	127341216	휘파람 친구	추수진 지음	샘터사	2019	9788946472983				813.8	1	1	2019-12-08
3	127341215	환상 동물 특	프로오르터	라이카미	2019	9791187504863				187	1	1	2019-12-08
4	127341214	하늘이 딱딱	신원미 지음	천개의바람	2019	9791190077187				813.8	1	1	2019-12-08
5	127341213	푸른 사자 와	이현 지음	오 창비	2019	9788936443054			305	808.9	1	1	2019-12-08
6	127341212	파피	애비 지음	원 보물창고	2019	9788961707305				843	1	1	2019-12-08
7	127341211	축구왕 이채	유우석 지음	창비	2019	9788936443061			306	808.9	1	1	2019-12-08

추출



# 1 데이터 수집 및 전처리

- 한국의 도서번호는  
국제표준번호(ISBN)에  
‘**부가기호**’를 덧붙여 구성  
(출처 : 서지정보유통원)

## • 부가기호표

→ 제1단계 : 독자대상기호 (제1행)

기호	대상독자	설명
0	교양	일반 독자층을 대상으로 한 것으로, 주로 전문적인 내용을 비전공 일반 독자들이
1	실용	- 주로 실무에 관계된 실용적인 내용의 도서 - 실생활에 활용할 수 있는 도서 - 일반인을 대상으로 한, 어떤 목적을 가진 수험서적
4	청소년	중·고등 학습 참고서에 해당되지 않는 것으로 중·고등학생을 대상으로 한 도서
5	학습참고서 1 (중·고교용)	중·고등학생을 대상으로 한 학습 참고서
6	학습참고서 2 (초등학생용)	초등학생을 대상으로 한 학습 참고서
7	아동	초등학습참고서에 해당되지 않는 것으로 영유아, 초등학생을 대상으로 한 도서
9	전문	주로 학술·전문적인 내용의 도서



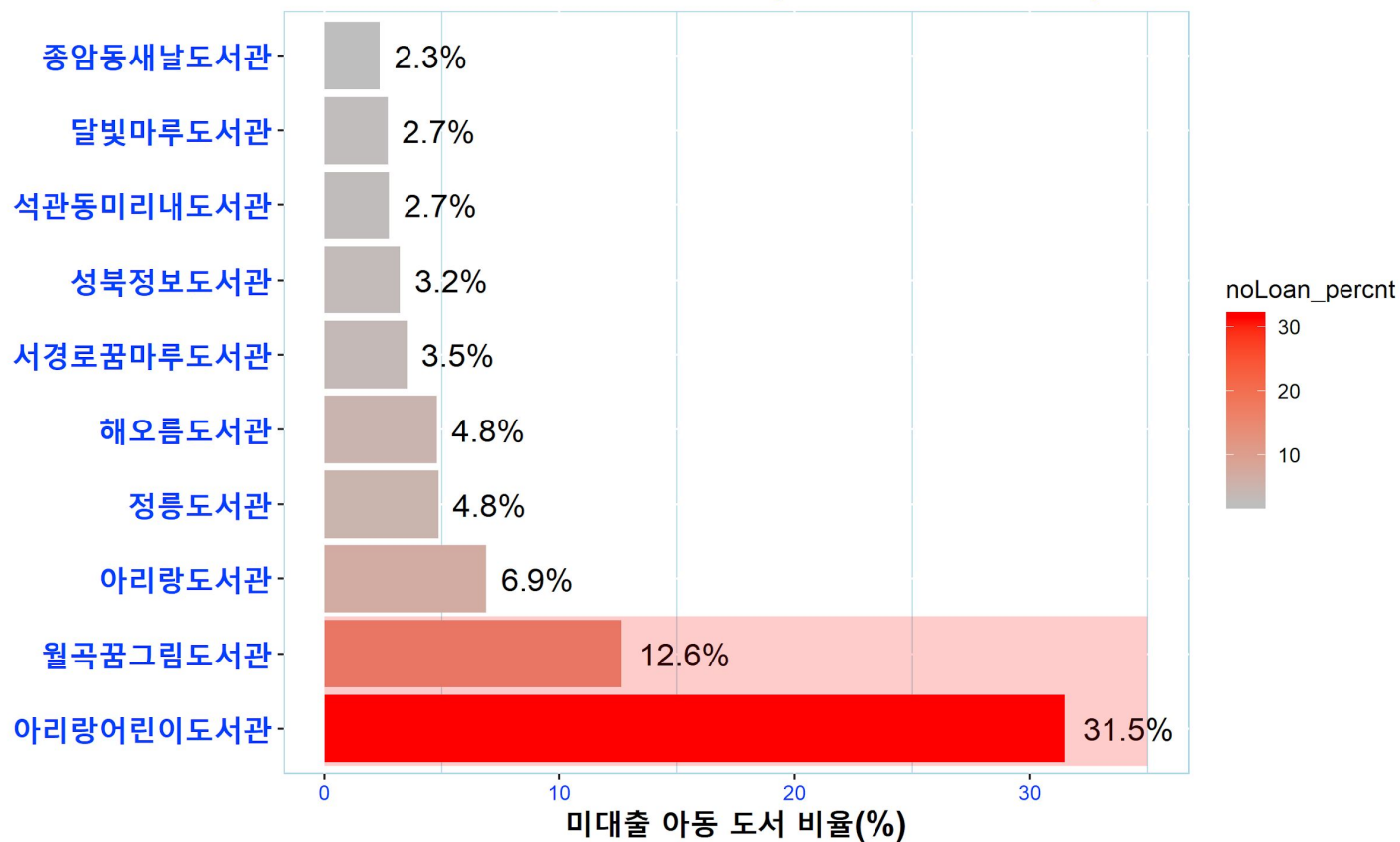
# 1 데이터 수집 및 전처리

도서관 별  
‘미대출 아동 도서 비율’,  
계산, 시각화

(누적 대출건수 0인 아동 도서 수)

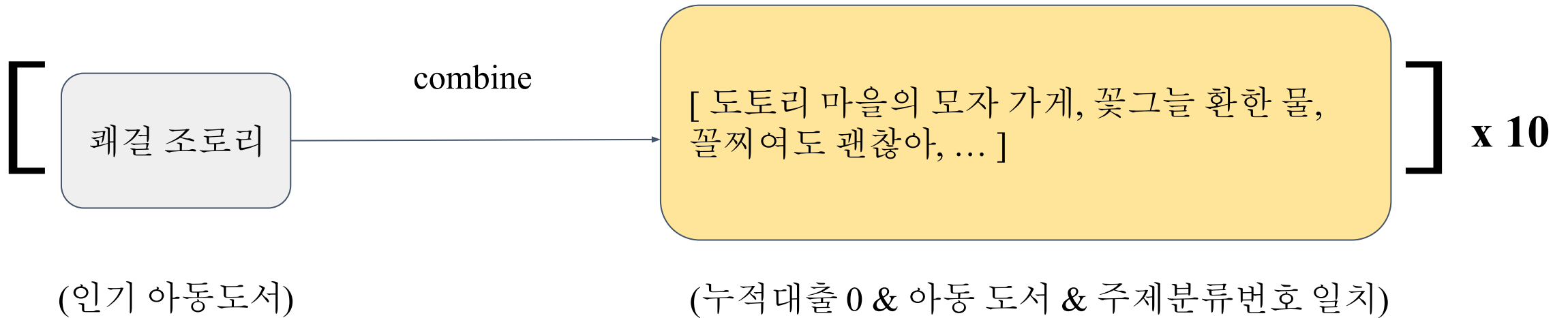
(전체 아동 도서 수)

2019년 12월 성북구 도서관별  
미대출 아동 도서 비율 (상위 10개 도서관)



# 1 데이터 수집 및 전처리

- '인기대출 도서 조회 API'를 활용해, 각 도서관의 최근 3개월 인기 아동도서 10개 수집
- '2019년 12월 장서/대출 데이터'에서, 인기 아동 도서와 주제분류번호 같은 도서 수집



# 1 데이터 수집 및 전처리

'도서 키워드 목록 API'를 활용,  
각 도서의 **키워드 & 가중치** 수집

※ 가중치 : 제목/목차/소개글에서  
단어의 출현 빈도

※ 가중치 값은 `sigmoid( )`  
함수로 스케일링 처리

```
> book_keywords_list[[1]]
```

```
$bookname
```

```
[1] "괘걸 조로리"
```

```
$ISBN
```

```
[1] "9788950925147"
```

```
$keywords
```

```
[1] "국내창작동화"
```

```
"김수정"
```

```
[3] "동화명작고전"
```

```
"을파소"
```

```
[5] "괘걸조로리19마법의성대결투" "하라유타카"
```

```
$weights
```

```
[1] 0.8807971 0.8807971 0.8807971 0.8807971 0.8807971 0.8807971
```

## 2 키워드 기반의 유사도 분석

인기 아동 도서에 대한 관련 미대출 장서(아동) 각각의 유사도 계산

(다차원 척도법 MDS 개념 활용)

키워드 별 가중치

미대출 장서

	김수정	그늘	플로리안	...
래걸 조로리	0.880797	0	0	
도토리 마을의 모자 가게	0	0	1	
꽃그늘 환한 물	0	0.55321	0	
꿀찌여도 괜찮아	0	0.12344	0	
...				

11행 \* 1500열의 행렬 (Bag of words) 일부

## 2 키워드 기반의 유사도 분석

	그림책	상상력	북뱅크	김숙
하늘 100층짜리 집	0.99	0.95	0.88	0.88
도깨비를 빨아버린 우리 엄마	1.00	0.95	0.00	0.00
도토리 마을의 모자 가게	1.00	0.88	0.00	0.00
돌돌 말아 김밥	0.88	0.88	0.00	0.00
왜냐면..	1.00	0.98	0.00	0.00
너는 어떤 씨앗이니?	1.00	0.00	0.00	0.00

전치행렬

	하늘 100층짜리 집	도깨비를 빨아버린 우리 엄마	도토리 마을의 모자 가게	돌돌 말아 김밥	왜냐면..	너는 어떤 씨앗이니?
그림책	0.99	1.00	1.00	0.88	1.00	1
상상력	0.95	0.95	0.88	0.88	0.98	0
북뱅크	0.88	0.00	0.00	0.00	0.00	0
김숙	0.88	0.00	0.00	0.00	0.00	0

## 2 키워드 기반의 유사도 분석

	김수정	그늘	플로리안	...
래걸 조로리	0.880797	0	0	

행렬곱  
(%\*%)

	래걸 조로리	도토리 마을의 모자 가게	꽃그늘 환한 물	꿀찌여도 괜찮아
김수정	0.8807971	0	0	0
그늘	0	0	0.55321	0.12344
플로리안	0	1	0	0
...				

‘래걸 조로리’(인기 도서)와 ‘도토리 마을의 모자 가게’(미대출 장서)간의 유사도(거리)  
계산

확장

( 1행 \* 1500열의 행렬 ) %\*% ( 1500행 \* 11열의 행렬 )

## 2 키워드 기반의 유사도 분석

	김수정	그늘	플로리안	...	내적값
괘걸 조로리	0.880797	0	0		4.1345
도토리 마을의 모자 가게	0	0	1		2.2345
꽃그늘 환한 물	0	0.55321	0		2.1101
꿀찌여도 괜찮아	0	0.12344	0		0.0023
...					

11행 \* (1500+1)열의 행렬

“내적값이 클수록, 인기도서와의 유사도가 크다”



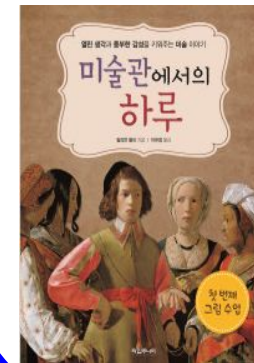
## 2 키워드 기반의 유사도 분석

	하늘 100 층짜 리 집	도깨 비를 빨아 버린 우리 엄마	돌돌 말아 김밥	왜나 면..	도토 리 마 을 의 모자 가게	너는 어떤 씨앗 이 니?	(Why? People) 정명훈	장수 탕 선 녀님	잠자 는 고 등 어	칫솔 맨 도 와 줘 요!	고맙 습니 다
아주아주	0.000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.00000000	0.9525741	0	0	0.00000000
랜드	0.000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.00000000	0.9525741	0	0	0.00000000
스파	0.000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.00000000	0.9525741	0	0	0.00000000
달샤베트	0.000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.00000000	0.9525741	0	0	0.00000000
게임방	0.000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.00000000	0.9525741	0	0	0.00000000
순간	0.000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.00000000	0.9525741	0	0	0.00000000
inner_products	36.17378	5.618169	5.139382	3.742309	3.257076	1.723694	1.3927750	0.9933063	0	0	0.00000000

유사도 기준 정렬

### 3 추천 도서 목록 선정

인기도서	추천도서	추천도서의 ISBN
pop book	recommend books	recommend ISBN
1 쾌걸 조로리	악당 반장!:김기정 동화집	9788990878823
2 쾌걸 조로리	힐링 썰매	9788953312074
3 쾌걸 조로리	미술관에서의 하루	9788991643437
4 쾌걸 조로리	꽃신:강소천 동화집	9788911027606
5 쾌걸 조로리	도깨비를 빨아버린 우리 엄마	9788970940199
6 쾌걸 조로리	도토리 마을의 모자 가게	9788901160924
7 쾌걸 조로리	한국을 이끈 사람들	9788943311117
8 쾌걸 조로리	노마의 발견:토론하는 위인 동화	9788954605137
9 쾌걸 조로리	전교 네명 머시기가 간다	9788954609623



## 사용 데이터

- 도서관별 인기대출 도서 API
- 도서관별 장서/대출 데이터  
(csv 파일)
- 도서 키워드 목록 API

Q & A

