

도서관 데이터 분석

: 독서 트렌드와 영향 요인 분석

데이터 사이언스 9조 데믹이들

강영민, 김성윤, 이성일, 이호건

목차 A table of contents.

1 프로젝트 목적 및 중요성

2 데이터 분석

3 가설 및 개선점

4 진행 상황

Part 1.

프로젝트 목적 및 중요성



프로젝트 목적 및 중요성:배경 · 목표 · 기대효과

도서관 데이터 가설
⇒ 도서관 데이터 분석
⇒ 독서 트렌드 파악
⇒ 현대인들의 생각 유추
⇒ 다양한 사업적 가치로 활용

배경

도서관 데이터를 통해서
독서 트렌드를 파악하고
영향 요인들을 파악

목표

독서 트렌드를 통한
도서관 프로그램 컨설팅

영향 요인을 통한
독서 트렌드 예측

독서 트렌드를 통한
현대인들의 심리 파악

기대효과

도서관 정보나루 사이트

- 2019~2022년까지 월별 1000건
- 나머지는 모두 ‘전체’ 선택

프로젝트 목적 및 중요성:데이터 선정

```
dt_1901 = pd.read_csv("BestLoanList_201901.csv", encoding='cp949')
dt_1901.sample(5)
```

| 순위 | | 서명 | 저자 | 출판사 | 출판년도 | 권 | ISBN | ISBN부가기호 | KDC | 대출건수 |
|-----|-----|--------------------------|-----------------------|----------|-----------|------|--------------|----------|-------|------|
| 978 | 977 | 그 겨울의 일주일 :메이브 빈치 장편소설 | 지은이: 메이브 빈치 ;옮긴이: 정연희 | 문학동네 | 2018 | NaN | 9.788950e+12 | 3840 | 843.5 | 564 |
| 425 | 426 | 내일은 실험왕 :본격 대결 과학실험 만화 | 글: 곰돌이 co. ;그림: 홍종현 | 대한교과서 | 2006-2008 | 4.0 | 9.788940e+12 | 77400 | 400.0 | 840 |
| 352 | 353 | Why? 소프트웨어와 코딩 | 글: 조영선 ;만화: 이영호 | 예림당 | 2015 | 64.0 | 9.788930e+12 | 74400 | 5.1 | 902 |
| 884 | 883 | 엉덩이 탐정 : 뽕뽕 사라진 과자를 찾아라! | 트를 (지은이), 전경아 (옮긴이) | 문학수첩 리틀북 | 2015 | NaN | 9.788960e+12 | 77830 | 833.8 | 588 |
| 880 | 879 | (빈대가죽의) 응답하라 짬돌이 시대 | 글: 스토리박스 ;그림: 류수형 | 재미박스 | 2016 | 28.0 | 9.788960e+12 | 77320 | 320.0 | 589 |

Part 2.

데이터 분석

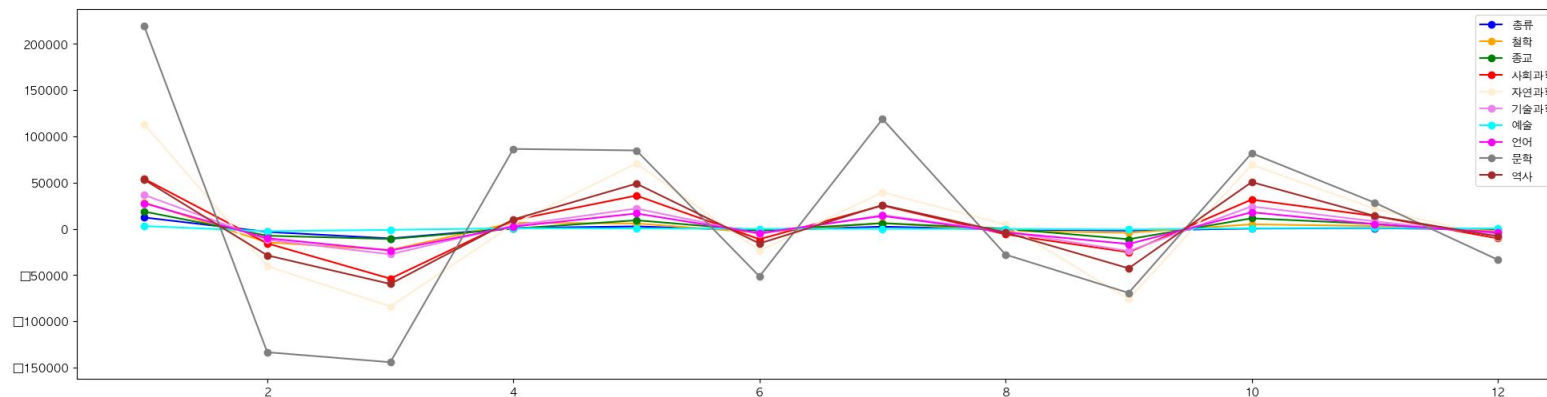


데이터 분석:분석 내용

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

급상승 주제 분석

- 2019~2022년 월별로 전월 대비 대출증가량이 가장 높은 주제 파악



데이터 분석:실험 결과

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

연도별 비교

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.

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

- 연도별 비교 결과, 1월에는 문학 주제의 대출 증가량이 높다는 특징 파악

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

데이터 분석:해석

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 성장률 논란

UNWTO “2019년 세계 여행객 15억명 돌파”

7월 :

한일 무역 분쟁

8월 :

영향 원인 파악

- 가설을 세워 다양한 데이터로 검증
- 사회적 이슈를 조사하여, 일반적인 경향을 따르지 않는 주제에 대해서 원인 파악

Part 3.

가설 및 개선점



가설

1. 독서 트렌드의 특이점과 사회적 이슈 사이의 상관관계가 있을 것이다.
2. 서점 데이터를 이용한다면 영향 요인에 대해 더 인과성을 설명할 수 있을 것이다.
3. 더 많은 연도를 분석한다면 일반적인 독서 패턴을 만들 수 있을 것이다.
4. 다양한 컬럼을 활용하여 분석한다면 일반적인 독서 패턴을 만들 수 있을 것이다.

개선점

1. 사회적 이슈를 어떻게 데이터화 할 것인가?
=> 크롤링을 통한 해당 월의 이슈를 데이터화
2. 독서 트렌드는 다양한 방면의 영향을 많이 받는다.
=> 서점 데이터를 활용하여 다양한 분석 시도

Part 4.

진행 상황



진행 상황:역할 분담

<데이터 선정>

- 도서관 정보나루 데이터 선정

1

<데이터 수집>

- 성윤 : 19년도 데이터 수집
- 호건 : 20년도 데이터 수집
- 영민 : 21년도 데이터

2

<데이터 분석>

- 모든 인원 : 데이터 분석 기법 선정
- 성윤 : 19년도 인기대출도서 데이터 분석
- 호건 : 20년도 인기대출도서 데이터 분석
- 영민 : 21년도 인기대출도서 데이터 분석

3

수집

<결과 해석>

4

- 성일 : 22년도 데이터
- 성윤 : 19년도 사회적 이슈 조사 및 연관성
- 호건 : 20년도 사회적 이슈 조사 및 연관성
- 영민 : 21년도 사회적 이슈 조사 및 연관성

파악

진행 상황: 일정 계획



중간 발표 전 :

데이터 선정, 간단한 데이터 분석을 통해서 결과 초안 제시



중간 발표 후 :

다양한 컬럼과 데이터를 활용하여, 독서 트렌드와 영향 요인 파악



Q&A

감사합니다