세미프로젝트

|  |
| --- |
| 서울시 도서관 별  대출 데이터 분석 |

**목 차**

**1. 프로젝트 개요**  **1-2**

1.1 프로젝트 기획 배경 및 목표 1

1.2 구성원 및 역할 1

1.3 프로젝트 추진 일정 2

**2. 프로젝트 결과**  **3-7**

2.1 데이터 수집 3

2.2 데이터 분석 3-6

2.3 데이터 분석 결과 6-7

**3. 기대 효과**  **7**

**1. 프로젝트 개요**

1.1 프로젝트 기획 배경 및 목표

1. 배경: R을 활용하여, 서울시 공공도서관 아동도서(ISBN 기준) 재배치 방안을 강구.

* 이를 통해 이용자들이 원하는 도서가 없어 번거로운 타관도서 대출을 신청하거나 대출을 포기하는 상황을 방지하고, 미대출 장서의 이용률을 제고하고자 함

1. 목표: 서울시 공공도서관 도서 대출 데이터 및 서울시 인구 데이터를 분석해, 도서관 별로 ‘다른 도서관에 보내야 할 도서 목록’ 도출

1.2 구성원 및 역할

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**2. 프로젝트 개발 결과**

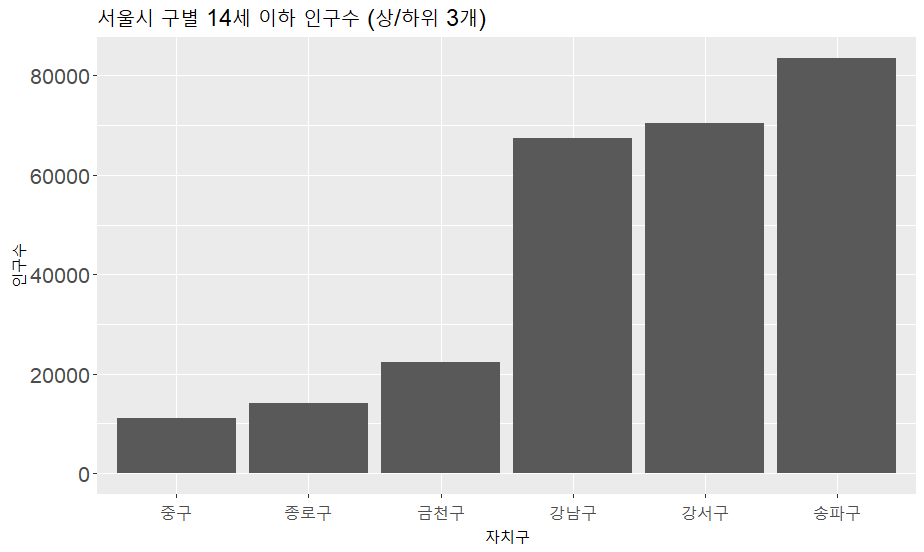
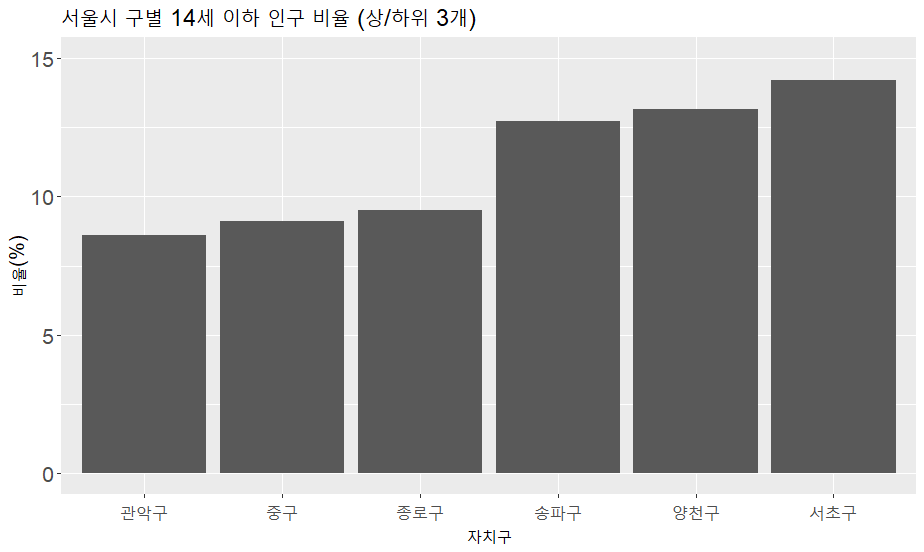
2.1 데이터 수집 (크롤링, API)

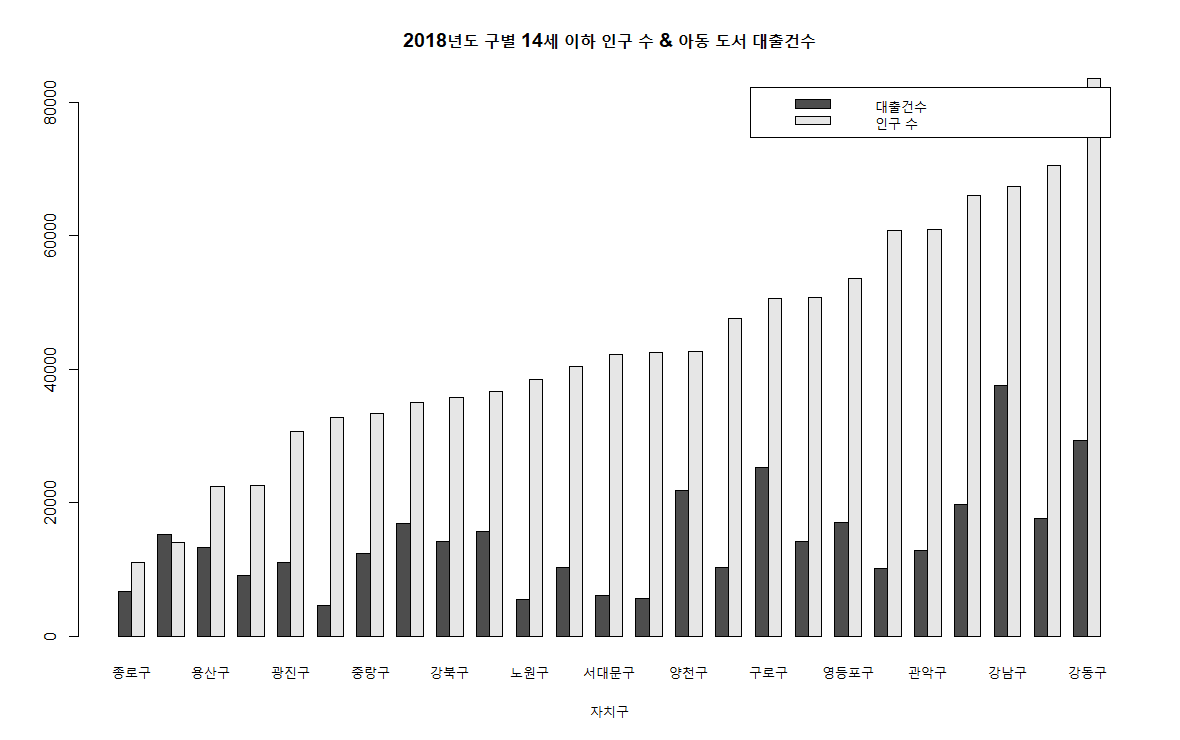
* 2016~2018년 서울시 구별 아동도서(ISBN 기준) 인기대출도서 데이터 (API 제공)
* 2018년 서울시 성북구/도봉구/강북구 도서관별 장서/대출 데이터 (csv 파일 크롤링)
* 서울시 구별 주민등록연앙인구(연령별) 통계 (csv 파일 다운로드)
* 서울시 구별 세대원수(세대원수별) 통계 (csv 파일 다운로드)

2.2 데이터 분석

2.2.1. EDA 및 시각화

* 2018년 서울시 구별 아동 인구비율 및 인구수 시각화 → 서울시 구별로 아동 인구비율 및 인구수가 상이하므로, 아동 관련 도서 대출 패턴도 상이할 것이라고 판단





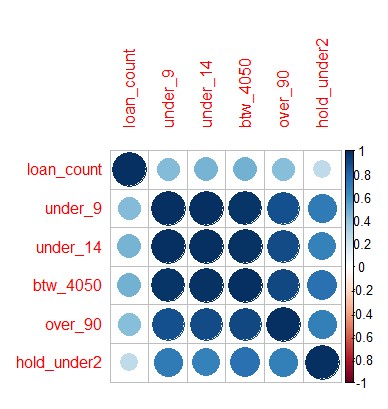
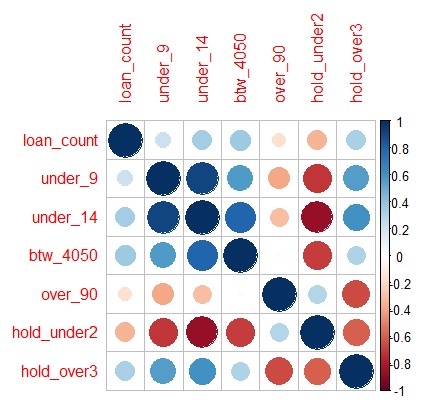
2.2.2. 상관관계 분석 및 시각화

2.2.2.1. 변수 설정

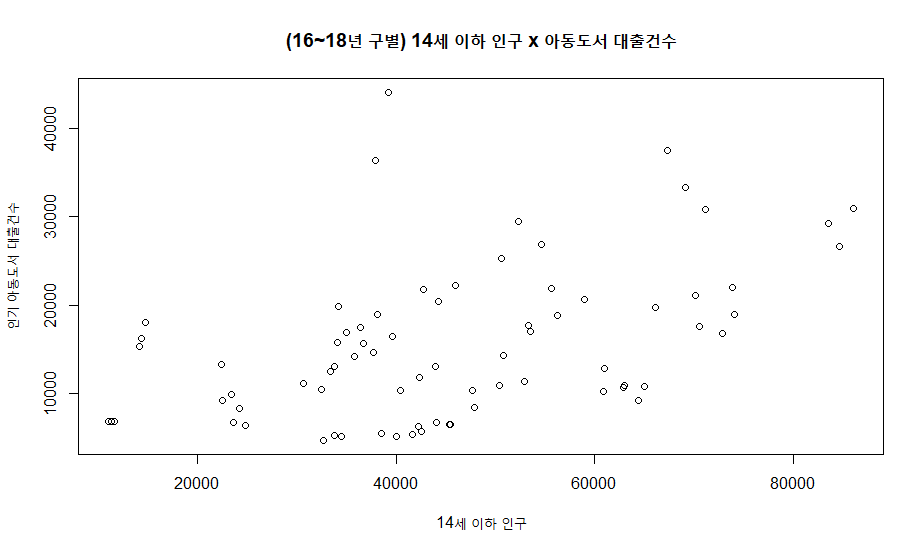
* <종속변수> 2016~2018년 서울시 구별 아동도서(ISBN 기준) 인기대출도서들의 총 대출 건수
* <독립변수 1> 서울시 구별 주민등록연앙인구(연령별) 중 9세 이하 인구 수 및 비율
* <독립변수 2> 서울시 구별 주민등록연앙인구(연령별) 중 14세 이하 인구 수 및 비율
* <독립변수 3> 서울시 구별 주민등록연앙인구(연령별) 중 40대 인구 수 및 비율
* <독립변수 4> 서울시 구별 주민등록연앙인구(연령별) 중 90세 이상 인구 수 및 비율
* <독립변수 5> 서울시 구별 세대원수(세대원수별) 중 2인 이하 가구 수 및 비율
* <독립변수 6> 서울시 구별 세대원수(세대원수별) 중 3인 이상 가구 비율

2.2.2.2. 결과 시각화

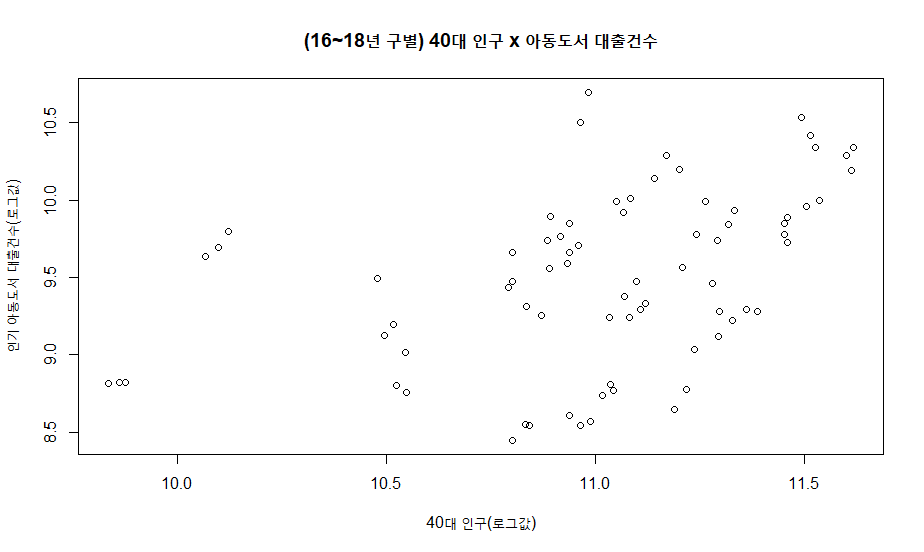
* 상관관계 그래프 시각화 ( 좌측 인구 절대수치, 우측 인구 상대비율 기준 )

* ‘14세 이하 인구 수’와 ‘아동 도서 대출 건수’ 간에 양의 상관관계 보임



* ‘40대 인구 수’와 ‘아동 도서 대출 건수’ 간에 양의 상관관계 보임



2.3 데이터 분석 결과 - 재배치 도서 선정

* 미대출 장서 비율 높은 도서관(둘리, 북서울중학교)에서, 송파구로 재배치할 도서 목록 도출
  + 기준 : 개관 이후 누적 대출건수 5건 미만이면서, 2018년 송파구 아동 인기도서 목록에 등재된 도서



* 14세 미만 인구 수가 적은 구(중구)에서, 송파구로 재배치할 도서 목록 도출
  + 기준 : 개관 이후 누적 대출건수 5건 미만이면서, 2018년 송파구 아동 인기도서 목록에 등재된 도서

****

**3. 기대 효과**

* 공공도서관 사용자 관점에서는 도서관 이용의 불편함(원하는 도서가 없어 번거로운 타관도서 대출을 신청하거나 대출 포기)을 줄이고, 운영자 관점에서는 미대출 장서의 이용률을 향상하며 타관도서 이동을 최적화할 수 있을 것으로 기대됨