

OUTLIER 최종 발표회

성범죄 예방을 위한 CCTV 설치 입지 선정
- 광진구를 중심으로

7팀 이재경(20201067)
김민영(20220823)

발표 순서

서론

1. 주제 선정 계기
2. 선행 연구 분석
3. 분석 방법

분석

1. 데이터 살펴보기
2. 시각화
3. 회귀분석

결론

1. 결론
2. 보완점

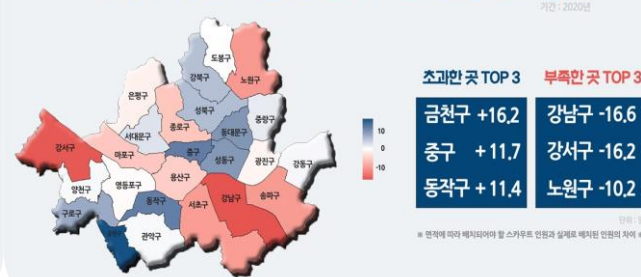
주제 선정 계기 - 문제현황1

• 서울시 정책 ‘여성안심 귀가서비스’의 효용성

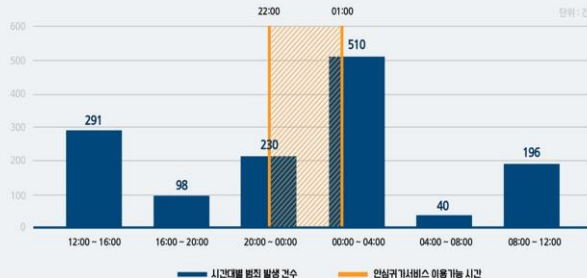
5개년 연도별 서울시 안심귀가서비스 이용실적과 5대 범죄, 강간 및 강제주행 발생건수



자치구 면적에 따른 스카우트 인원 배치 비교



시간대별 성폭력 범죄 발생 건수 (2018년)



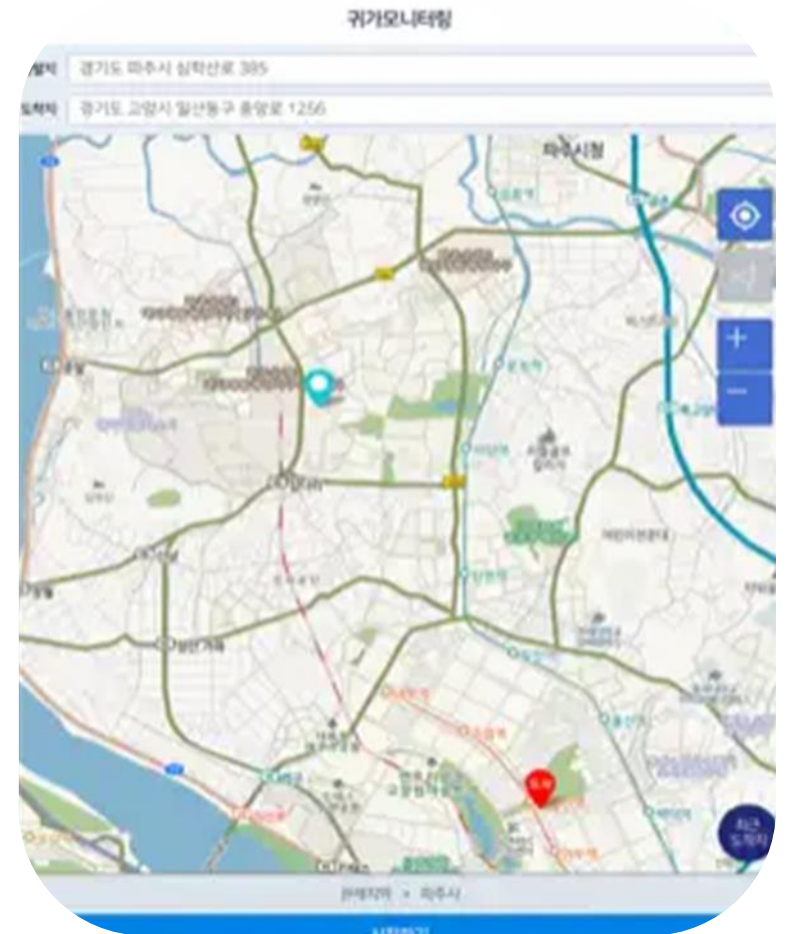
요일별 강력 범죄 발생 건수



- 늘어나는 서비스 이용률에 비해 성범죄 발생률은 변화 없음
- 스카우트 배치가 수요에 비해 원활히 이루어지지 않고 있음
- 서비스 운영시간이 평일 22시-01시로 상시 성범죄를 예방하기 힘들

주제 선정 계기 – 문제현황2

- 서울 안심이 앱의 효용성
 - 서울 안심이 앱: 여성이 안전하게 귀가할 수 있도록 주변 CCTV를 모니터링 해주는 서비스
 - 안심이 앱의 낮은 이용률 - 2021년 5월 평균 239건으로 전년도 대비 200여건 줄어듦
 - 안심이 앱의 신뢰성 문제
 - 가장 큰 문제는 ‘CCTV가 없는 귀갓길이라면 안심이 앱도 소용 없음’
 - 따라서 안심이 앱의 신뢰성을 높이기 위해서라도 CCTV 사각지대를 없애는 것이 급선무



주제 선정 계기 - 해결방안



- CCTV를 통한 상시 감시로 성범죄 예방
 - 데이터를 기반으로 범죄에 취약한 구역을 파악해 취약 구역에 CCTV설치를 제안

주제 선정 계기 - 해결방안

성범죄가 빈번한 것과 깊은 연관이 있는 **요소**를 파악

회귀분석으로 요소와 범죄지수간 **긴밀성** 숫자화

긴밀성을 기반으로 CCTV 필요성을 점수 매겨 필요성이 큰 지역 선정

선정된 구역 중 CCTV 사각지대를 찾아 추가 설치 제안

선행 연구 분석

2.2.4 지수 및 세부지표 산출방법

■ 최종 우선설치지수 산출식은 아래와 같으며 최대 100점 만점으로 산출됨

| 그림 2-10 | 최종 우선설치지수 산출식

$$\text{최종 우선설치지수} (\leq 100\text{점}) = \sum (\text{세부지표별 점수} \times \text{세부지표별 가중치})$$

■ 최종 우선설치지수를 구성하는 각 세부지표별 정의는 아래와 같음

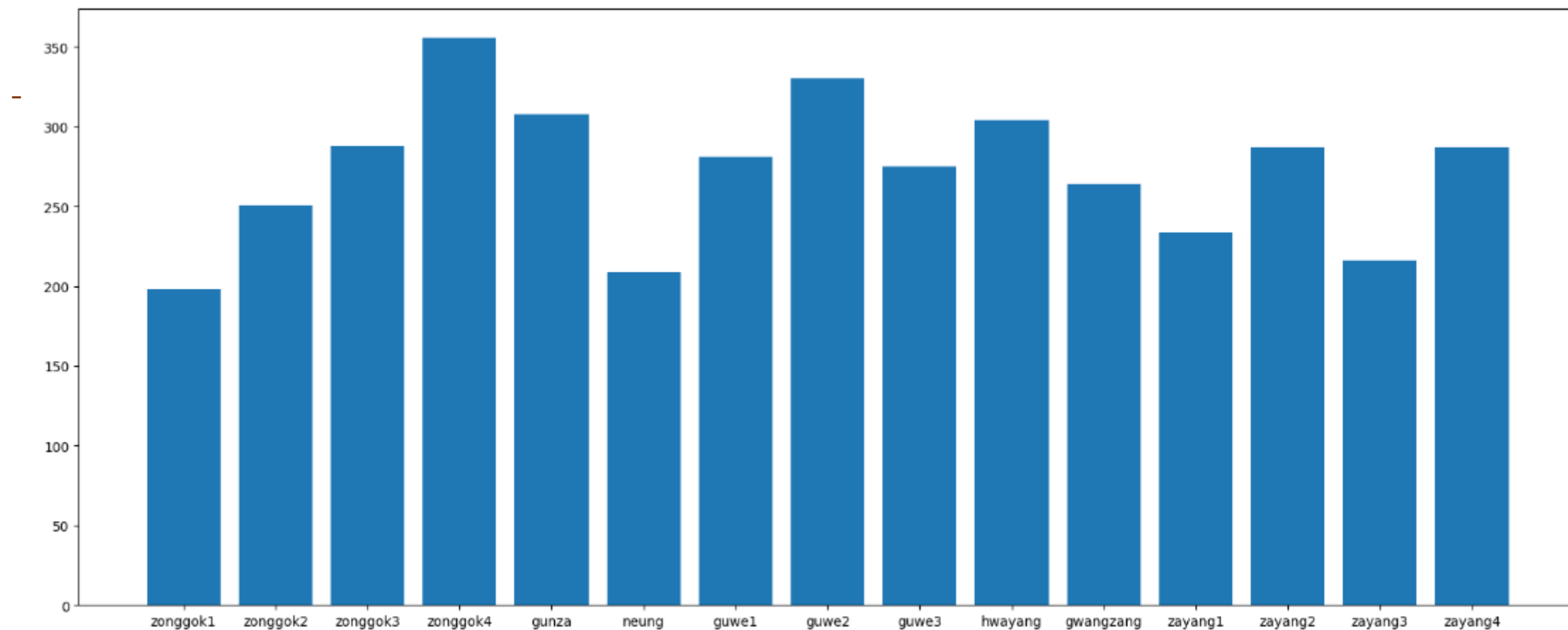
| 표 2-5 | 세부지표 정의

연번	세부지표	정의
1	감시취약지수	기설치된 공공 CCTV가 감시하지 못하는 취약정도
2	범죄발생지수	5대 범죄 실제 발생건수 기반 과거 범죄발생정도
3	풍속업소지수	유흥업소, 숙박업소 밀집정도(전체업종 대비)
4	유동인구지수	유동인구 밀집정도(전체시간대, 전연령)
5	부양인구지수	범죄취약인구(유소년인구, 고령인구) 밀집정도
6	외국인위험지수	생산가능인구 대비 등록 외국인 밀집정도
7	건물특성지수	건물 특성*에 의한 잠재 위험정도 *무벽건물/가건물, 일반주택/연립주택, 아파트, 주택외건물
8	아동시설지수	아동 및 청소년 교육시설 인근 설치 필요정도
9	공적감시취약지수	파출소, 지구대 등 공적감시시설이 감시하지 못하는 취약정도
10	대체감시취약지수	보안등과 같은 방범시설이 감시하지 못하는 취약정도

- 서울시를 격자 단위로 나눠 분석함
- 세부지표를 정의해 가중치를 부여한 뒤 최종 우선설치지수를 산출함
- 최종 우선설치지수가 높은 부지를 CCTV 설치 입지로 선정

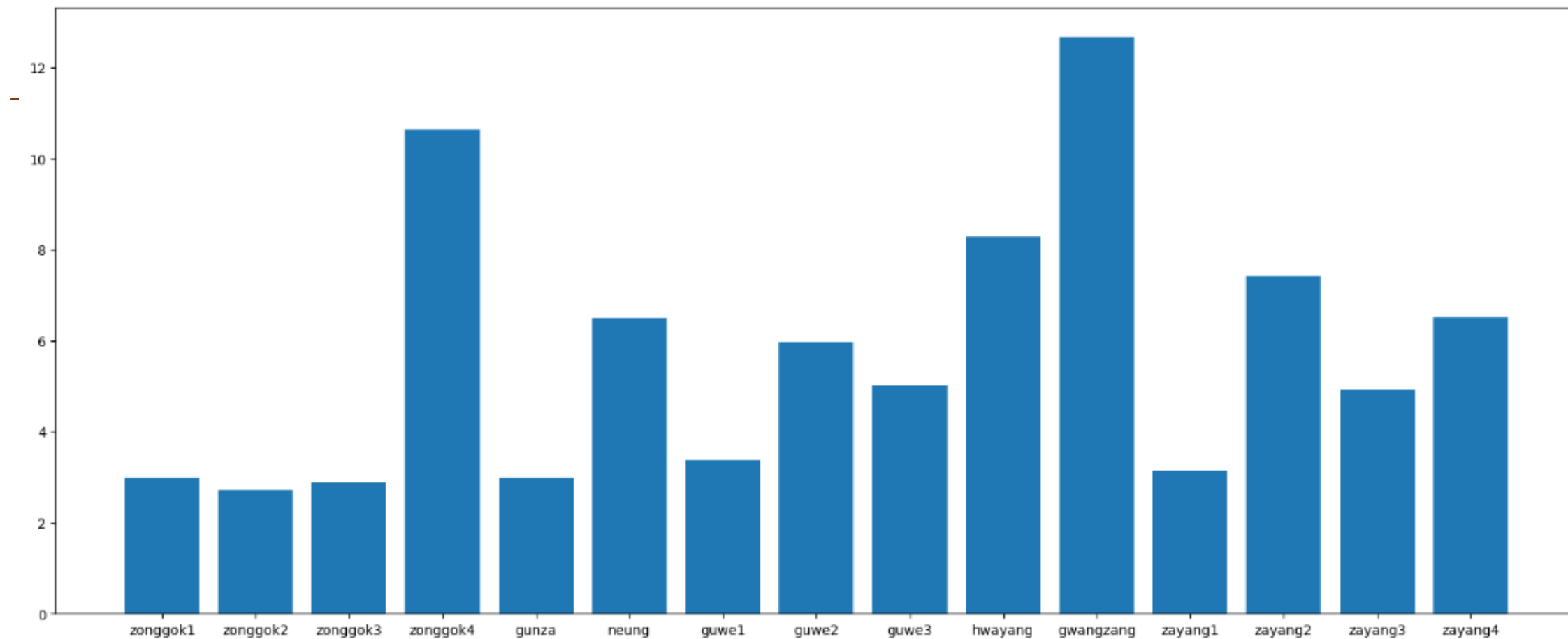
데이터 살펴보기

- 동 별 CCTV 개수



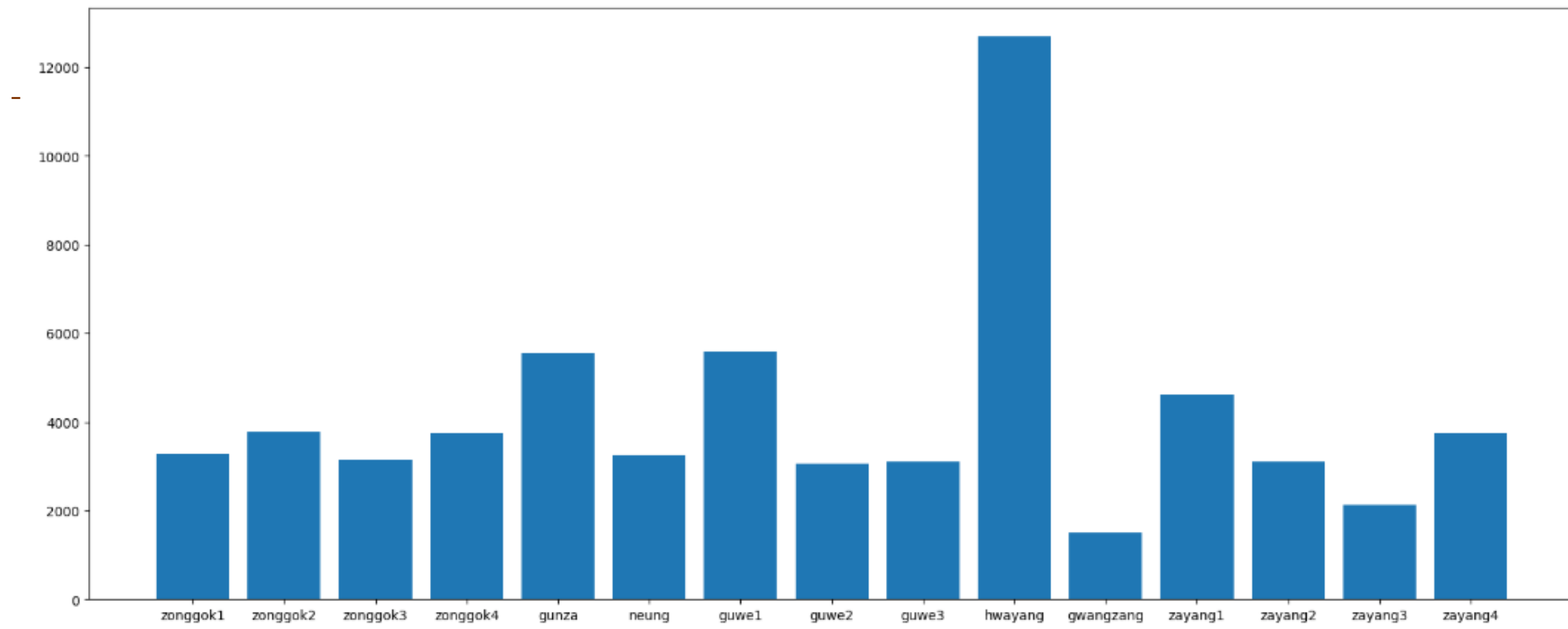
데이터 살펴보기

- 여성인구밀도



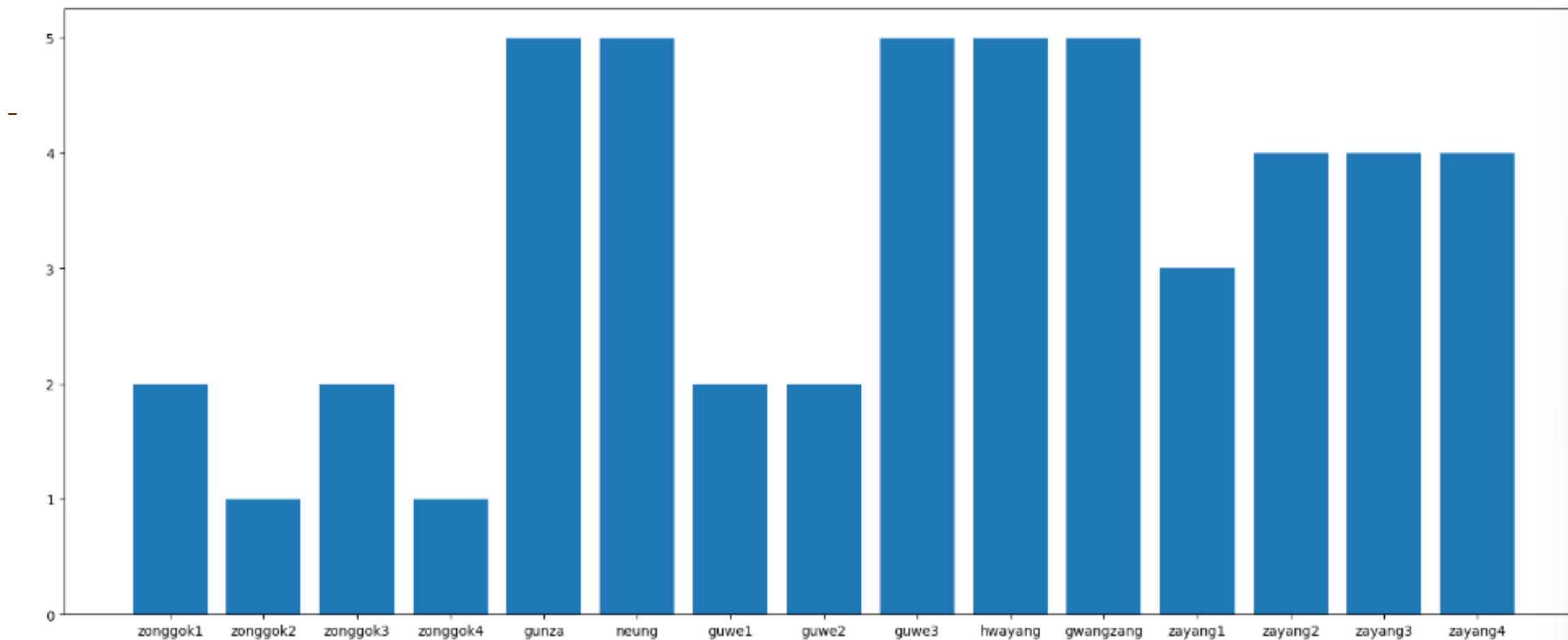
데이터 살펴보기

- 1인가구 수



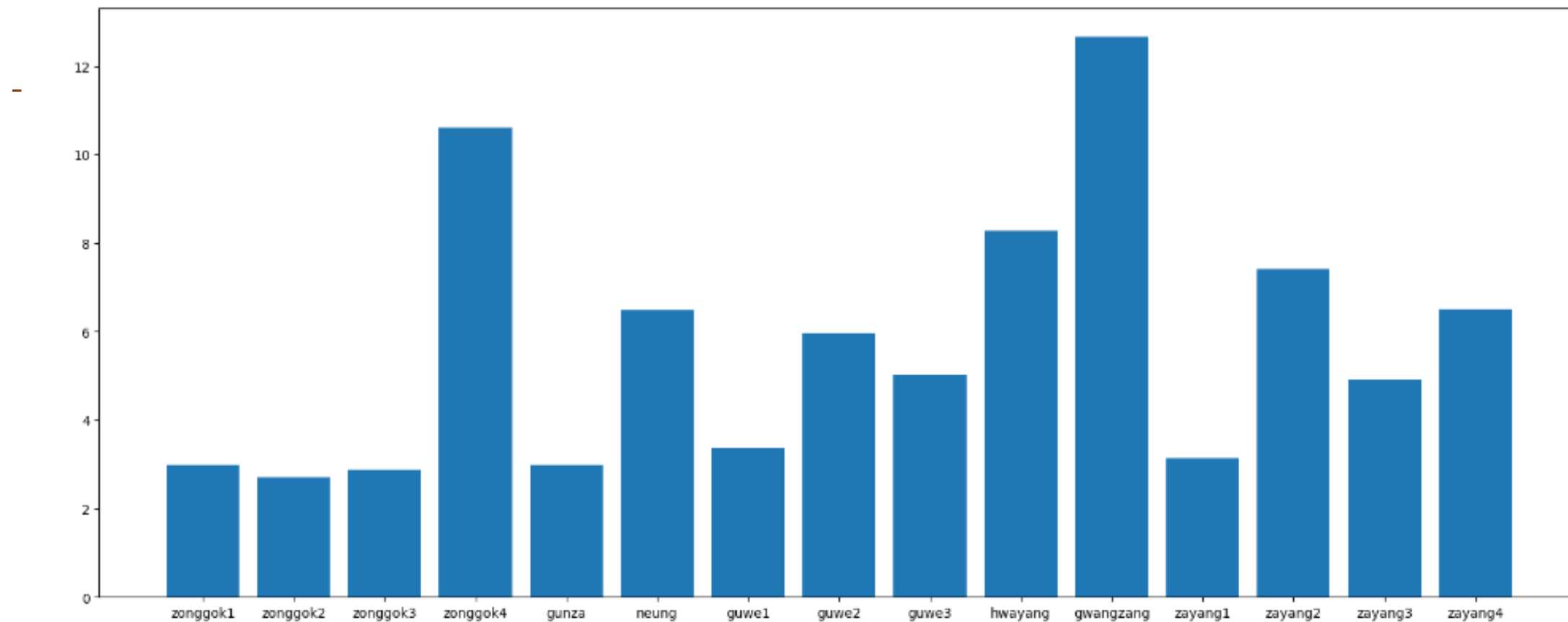
데이터 살펴보기

- 성범죄지수



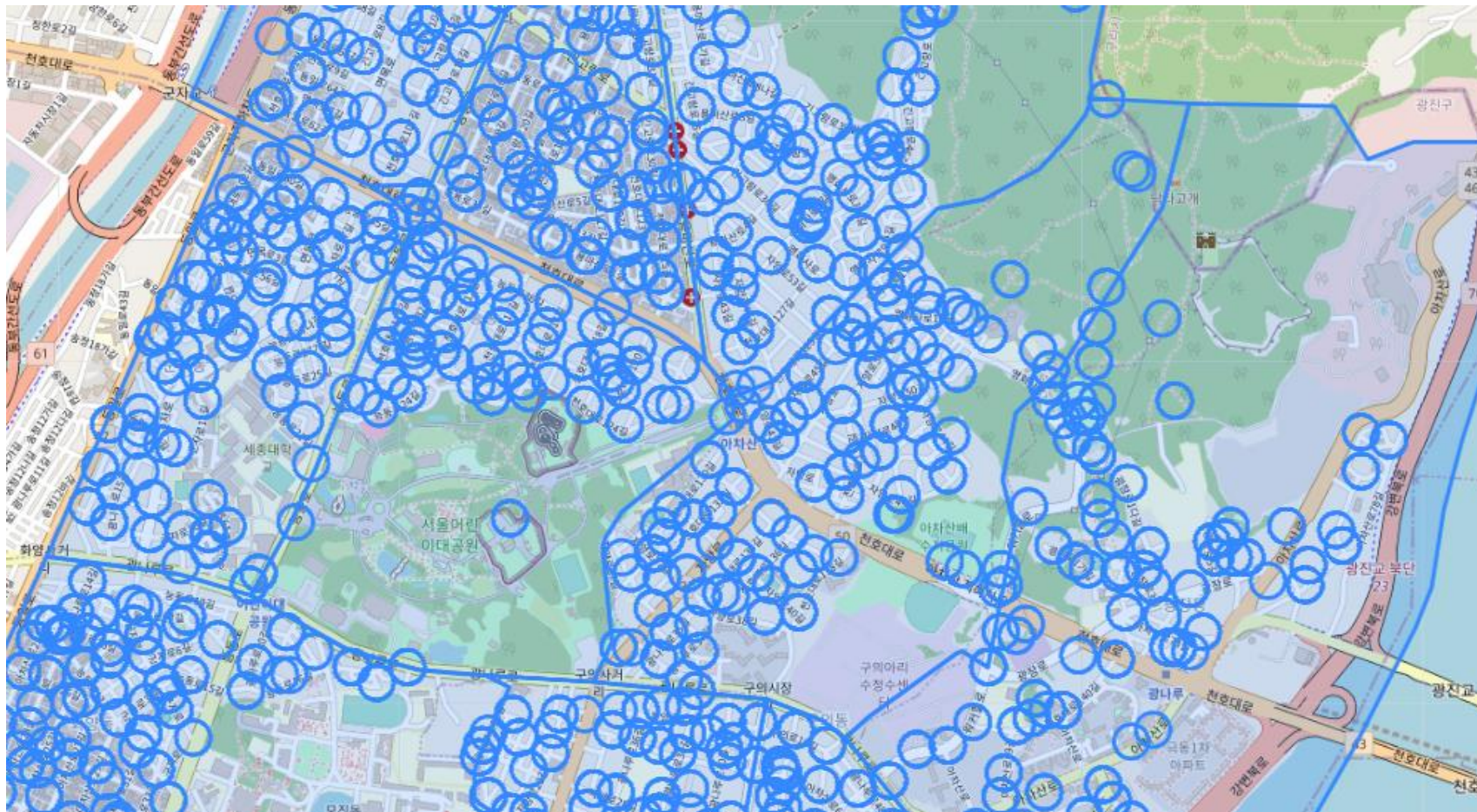
데이터 살펴보기

- 여성유동인구수



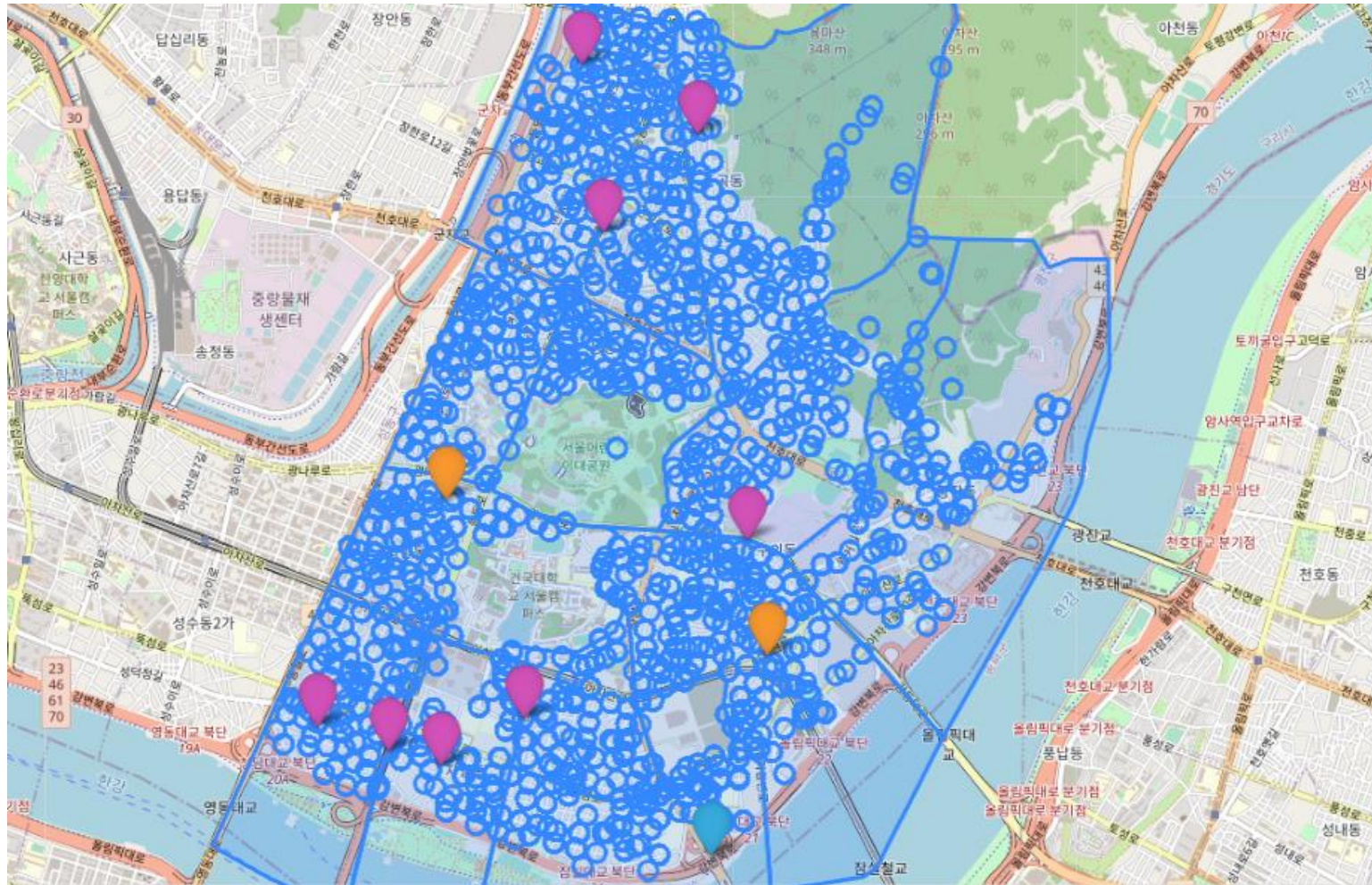
지도 시각화

- CCTV 유효거리가 50m 인 점을 반영해 CCTV 반경 지도시각화



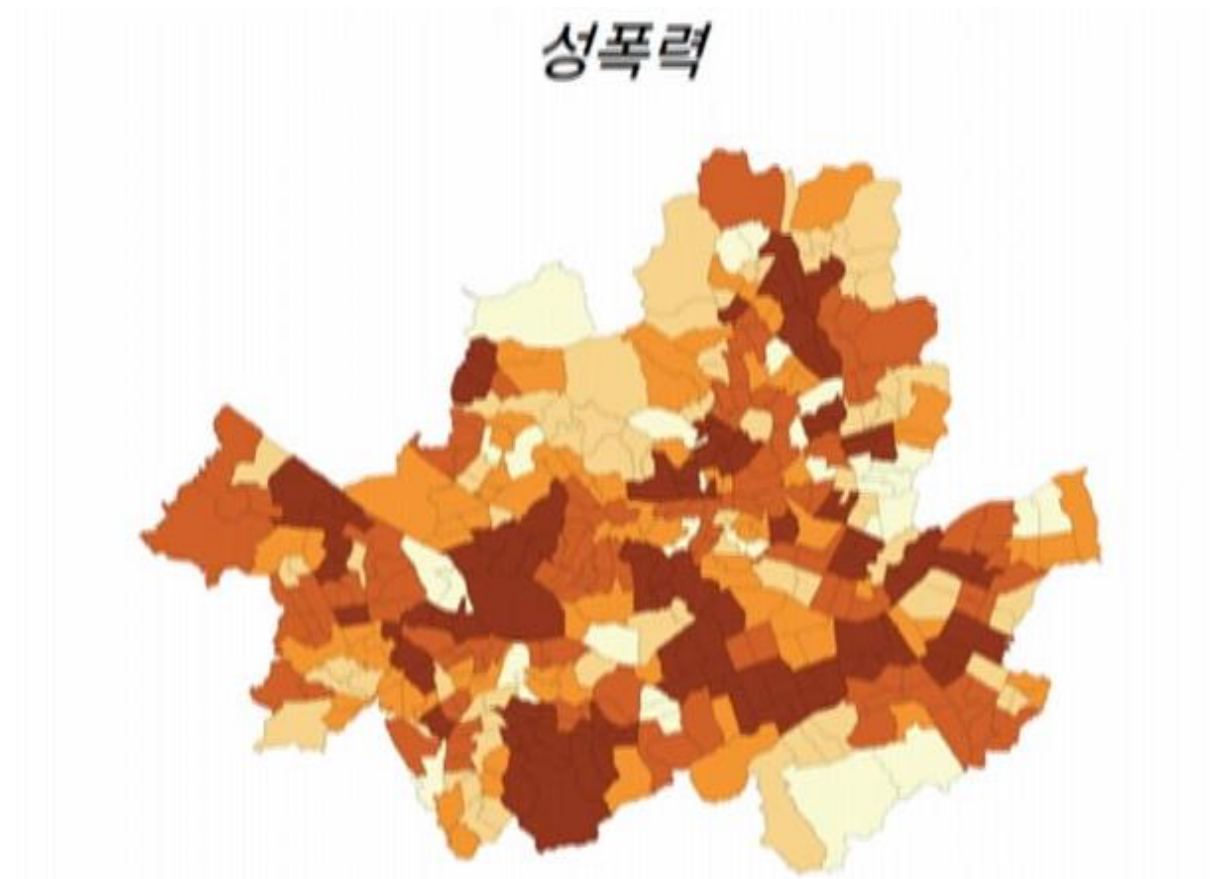
지도 시각화

- 경찰서, 파출소, 지구대 위치 지도에 표시



다중회귀분석

- 반응변수: 해당 동의 성범죄지수
- 설명변수:
 - (1) 동 규모 대비 1인가구 비율
 - (2) 동 규모 대비 여성 유동인구 비율
 - (3) 여성 거주인구 비율
 - (4) 유흥업소 개수
 - (5) CCTV 개수
 - (6) 평균 안심택배 서비스 이용률
 - (7) 인구밀도
 - (8) 지구대/파출소 유무



다중회귀분석

- 8개의 변수를 회귀분석에 사용할 시 다중공선성* 문제 있었음 -> 따라서 변수 개수를 줄여서 분석

	features	VIF
0	cctv	46.4
1	ansimdelivery_evg	4.4
2	woman_density	4226906.8
3	police_is	4.5
4	pop_density	13.7
5	single_household	10.8
6	pub	4.5
7	flo_pop	4227835.9



	features	VIF
0	ansimdelivery_evg	4.3
1	police_is	4.0
2	pop_density	12.3
3	cctv	44.8
4	single_household	9.5
5	pub	4.0
6	flo_pop	13.7

* 다중공선성: 설명 변수 중에 서로 상관성이 높은 것이 포함되어 있을 때는 분산·공분산 행렬의 행렬식이 0에 가까운 값이 되어 회귀 계수의 추정 정확도가 매우 나빠지는 일이 발생하는데 이 현상을 다중공선성이라 함.

다중회귀분석

- 회귀분석 결과

features	coef
ansimdelivery_avg	0.0
police_is	2.5
pop_density	-0.0
cctv	-0.0
single_household	0.0
pub	1.2
flo_pop	0.0

- 유의미한 회귀계수 결과를 얻지 못했음

: 성폭력 지수를 파악하기에 유의한 변수를 찾지 못했거나,
변수/데이터가 부족하다는 것의 문제점 때문으로 예상됨

* coef(coefficient): 데이터로부터 얻은 계수의 추정치

결론

- 유의미한 회귀계수를 얻지 못했으므로, 범위에 취약한 세부적인 지역 선정하는 대신 임의로 CCTV가 필요한 행정동을 선정
 - 기준은 점수 0에서 시작하여
 - (1) 여성 야간유동인구수가 평균보다 높을 경우 +1
 - (2) 성범죄지수가 가장 높은 5등급일 경우 +1
 - (3) 동 별 CCTV 개수가 평균보다 낮을 경우 +1 하여
- > 가장 높은 점수를 가진 행정동을 CCTV 필요 동으로 선정

결론

- CCTV 개수가 평균보다 낮은 동 추출

```
cctv_mean = df['cctv'].mean()
df[df['cctv'] < cctv_mean].index

Index(['zonggok1', 'zonggok2', 'neung', 'gwangzang', 'zayang1', 'zayang3'], dtype='object', name='dong')
```

- 여성유동인구 수가 평균보다 높은 동 추출

```
flo_mean = df['flo_pop'].mean()
df[df['flo_pop'] > flo_mean].index

Index(['zonggok4', 'neung', 'guwe2', 'hwayang', 'gwangzang', 'zayang2',
      'zayang4'],
      dtype='object', name='dong')
```

결론

- 성범죄지수 5등급인 동 추출

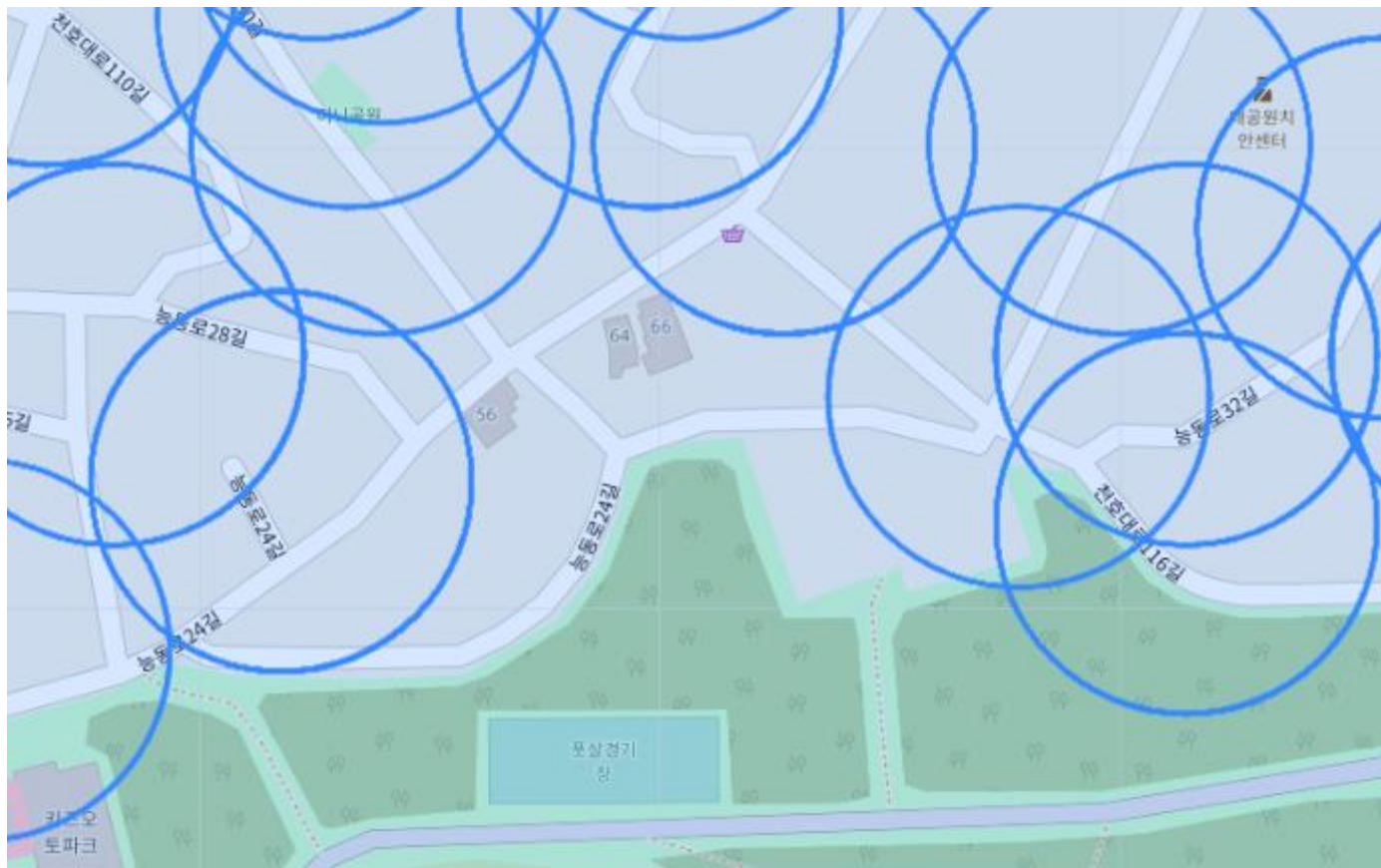
```
df[df['sexualviolence_grade'] == 5].index  
Index(['gunza', 'neung', 'guwe3', 'hwayang', 'gwangzang'], dtype='object', name='dong')
```

- 임의로 선정한 CCTV 필요 점수가 3점인 동은 총 두 개

: 능동, 광장동

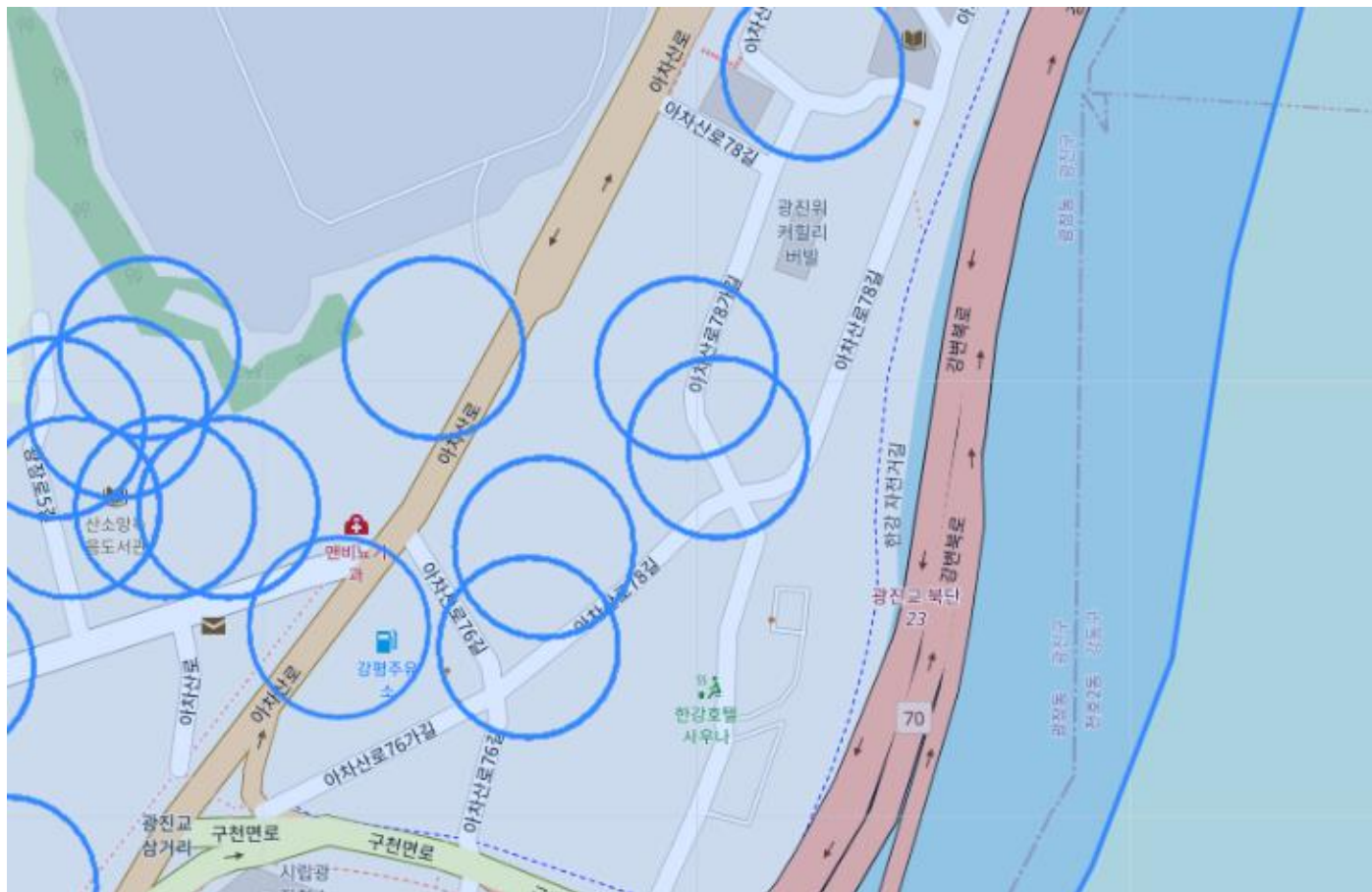
따라서 능동, 광장동의 CCTV 분포지역 중 사각지대를 찾아 CCTV 설치 제안

결론



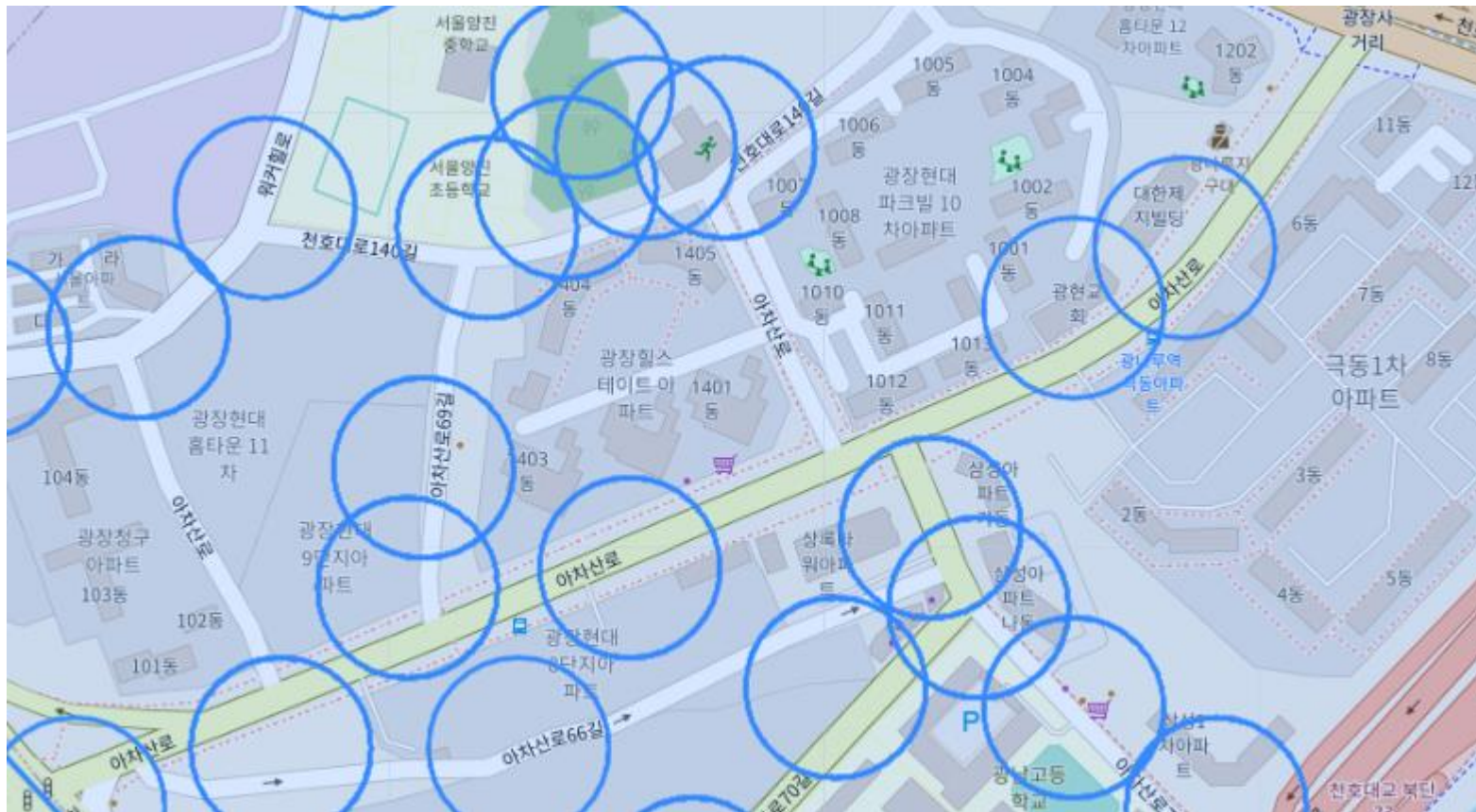
- 능동로 24길

결론



- 아차산로 78길

결론



- 아차산로 광장힐스테이트 아파트 남쪽

보완점

적절한 변수를 찾지 못해 회귀분석에 실패

보완을 위해서는 범죄를 설명할 수 있는 변수를 많이 찾는 것뿐만 아니라 범죄 발생과 변수가 연관이 될 것이라고 막연히 생각하여 데이터를 수집하는 것이 아닌, 범죄를 발생시키는 데 확실한 연관이 있다는 근거가 마련된 변수를 수집하는 것이 최우선으로 보여짐

반응변수 표본 확보의 어려움

반응변수를 해당 동의 성범죄지수로 설정했는데, 보안상의 이유로 성범죄지수가 제공되지 않은 행정동도 존재했음. 이로 인해 반응변수가 0인 관측값도 존재하게 되어 분석 결과에 영향이 갔을 것으로 보임. 반응변수를 범죄지수가 아니라 그에 대체할 다른 변수를 찾아서 다시 분석할 필요가 있음.

**발표를 들어주셔서
감사합니다**