

+

공공 데이터를 활용한 서울특별시 ‘여성 1인 가구’ 안전성 확인 및 해결 방안 탐색

프로젝트 기간 : 2022.05.18 ~ 06.09

프로젝트 참여 인원 : 김해인, 조가운, 주은정

+

목차

01

- 팀소개
- 분석 배경 및 목적

02

- 분석 과정 및 분석 기법
- 상관분석, 회귀분석, Kmeans 군집화

03

- 분석결과
- 시각화, 위험지역 순위화

04

- 시사점 및 개선점
- 참조 자료

팀 소개



팀장 : 김해인
데이터 수집
PYTHON 분석



팀원 : 조가운
데이터 변환
데이터 시각화



팀원 : 주은정
데이터 전처리
R 분석

분석 배경 및 목적

서울시 1인 가구 증가에 따른 연구 필요성 대두

서울시의 1인가구 수는 2000년 50만명에서부터 2015년 100만명을 넘기고, 2019년에는 약 130만명에 이를만큼 꾸준히 증가하고 있다. 또한 주된 가구 구성방식이 핵가족 단위에서 1인가구 단위로 바뀌어가고 있다. 학계에서도 많은 연구가 이루어지고 있기에 지자체에서도 이에 주목하고, 구체적 정책을 제시해야 한다. 서울특별시 1인가구 대책정책 연구(서울시, 2014)에서는 1인가구 정책을 '주거주택 부문', '사회적안전망 부문', '공유사회 부문' 세 가지로 나누고 있는데, 본 연구에서는 '사회적안전망 부문'에 주목해 보았다.

여성 1인 가구에 대한 정책 지원 필요

1인가구 밀집지역은 비밀집지역에 비해 2~3배 높은 범죄발생율을 보인다. 또한 주요 5대 강력범죄가 주거지에서 가장 많이 발생하며, 여성 피해자의 수도 꾸준히 증가세를 보이고 있다. 빅카인즈를 통해 2019년 5월 '신림동 감간미수 사건' 전후로 '1인가구 and 여성' 키워드로 연관어 분석을 해보았을 때, 주거공간에서의 범죄불안 및 지원정책에 대한 수요가 증가했음을 알 수 있었다. 따라서 본 연구에서는 정책지원 필요성이 크게 증가했음을 인지하고, 이에 대해 구체적인 지원이 이루어져야 하기에 여성 1인가구에 대한 정책을 찾아 분석했다.

+

분석 배경 및 목적

1인 가구 증가

여성 1인 가구의 안전
문제 대두

=

1인 가구 여성 안전 정책 필요!

낮은 정책 인지도
및 범죄율 증가

?

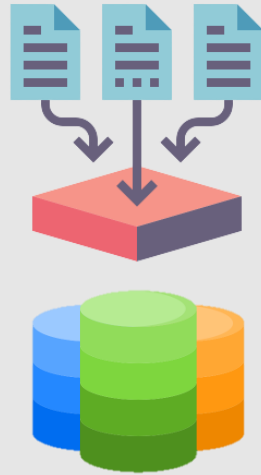
+

분석 순서

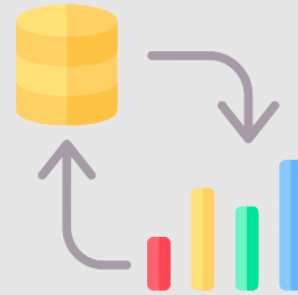
데이터셋 선택



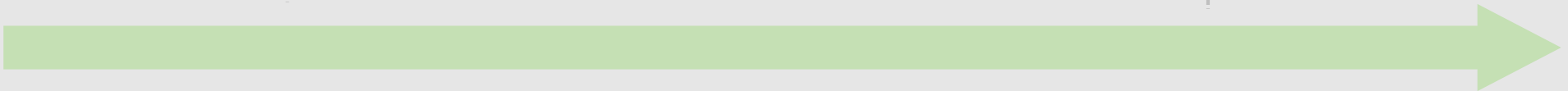
데이터 전처리



데이터 변환

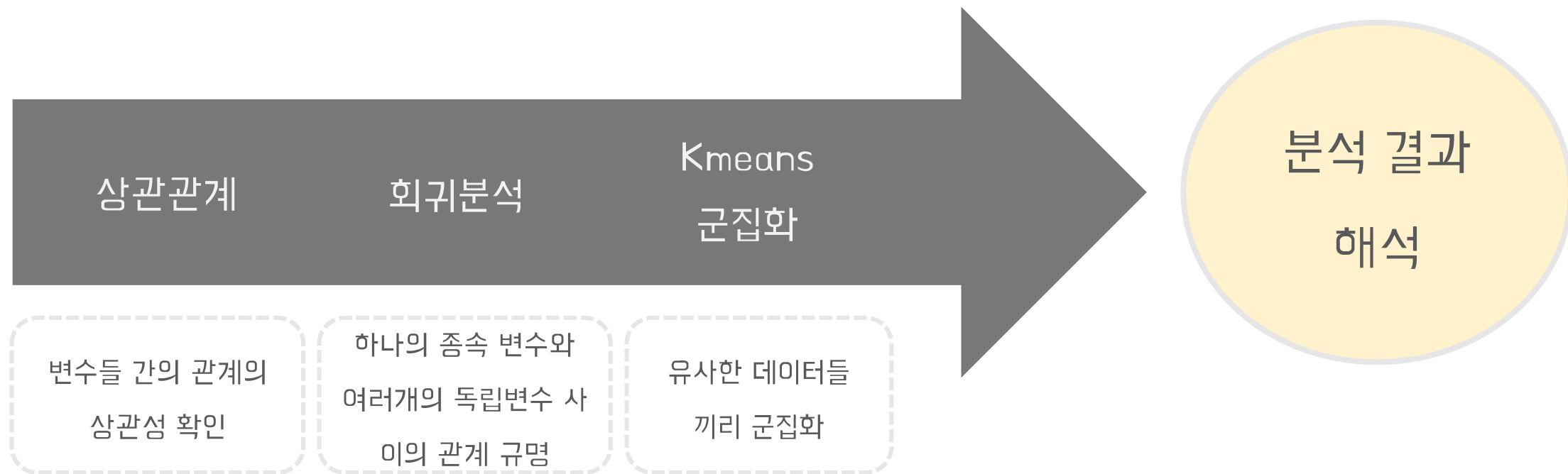


데이터 마이닝
/ 데이터 시각화



+

분석 과정



+

+

분석 과정

상관분석

회귀분석

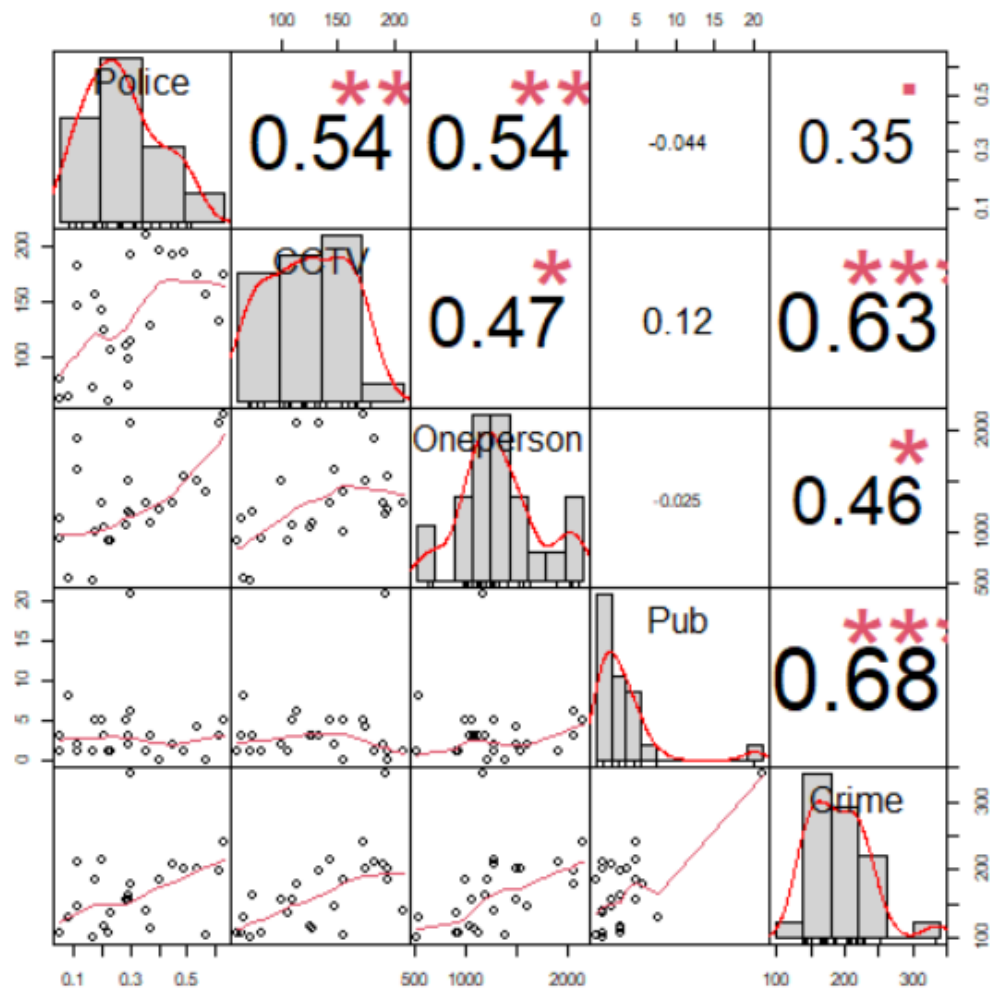
Kmeans 군집분석

위험지역 순위화

+

+

분석 과정 상관분석



0.68 crime - pub

0.63 crime - cctv

0.54 police - cctv

0.54 police - oneperson

+

분석 과정 회귀분석

```
Call:
lm(formula = df$crime ~ df$police + df$cctv + df$one_person +
    df$pub, data = df)
```

Residuals:

Min	1Q	Median	3Q	Max
-46.068	-12.757	7.566	17.763	32.122

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	22.01525	18.86228	1.167	0.25687
df\$police	0.92130	38.77338	0.024	0.98128
df\$cctv	0.50120	0.13587	3.689	0.00145 **
df\$one_person	0.03565	0.01481	2.408	0.02582 *
df\$pub	8.44544	1.26010	6.702	1.6e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 25.36 on 20 degrees of freedom
Multiple R-squared: 0.8271, Adjusted R-squared: 0.7926
F-statistic: 23.92 on 4 and 20 DF, p-value: 2.21e-07

유의한 변수 :

cctv, one_person, pub

유의하지 않은 변수 :

police

회귀분석을 통해 police의 p-value 값이
0.05보다 크기 때문에 유의하지 않음을
확인할 수 있다.

Kmeans 군집화 1차 군집분석

각 군집 별 개수

```
> clustering_km$size
[1] 1 13 11
```

1차 군집분석 중심점

```
> clustering_km$centers
      cctv one_person      pub      crime
1  1.2182396 -0.2699558  4.1968837  3.1757632
2  0.6902045  0.6229750 -0.1886902  0.3556567
3 -0.9264453 -0.7117018 -0.1585373 -0.7090274
```

2 군집 :

CCTV비율이 2번째로 높지만 3군집에 비해
범죄율과 여성 1인가구의 비율이 가장 높음

1 CLUSTER	2 CLUSTER	3 CLUSTER
중구	동대문구	마포구
	동작구	송파구
	관악구	강동구
	광진구	강서구
	금천구	용산구
	중랑구	은평구
	영등포구	강북구
	서대문구	노원구
	구로구	도봉구
	양천구	종로구
	강남구	서초구
	성북구	
	성동구	

Kmeans 군집화 2차 군집분석

2차 군집분석 중심점

```
> clustering_km$centers
      cctv one_person      pub      crime
1  1.2182396 -0.2699558  4.1968837  3.1757632
2  0.6902045  0.6229750 -0.1886902  0.3556567
3 -0.9264453 -0.7117018 -0.1585373 -0.7090274
```

범죄율이 가장 높고,
1인가구수, 유흥주점이 가장 많으며,
cctv 설치가 부족한 지역



동대문구
동작구
관악구

+

분석 결과

시각화

위험지역 순위화

+

분석 결과 범죄발생 건수와 유의미한 변수



범죄물

: crime



유흥주점

: pub



여성 1인가구

: one_person

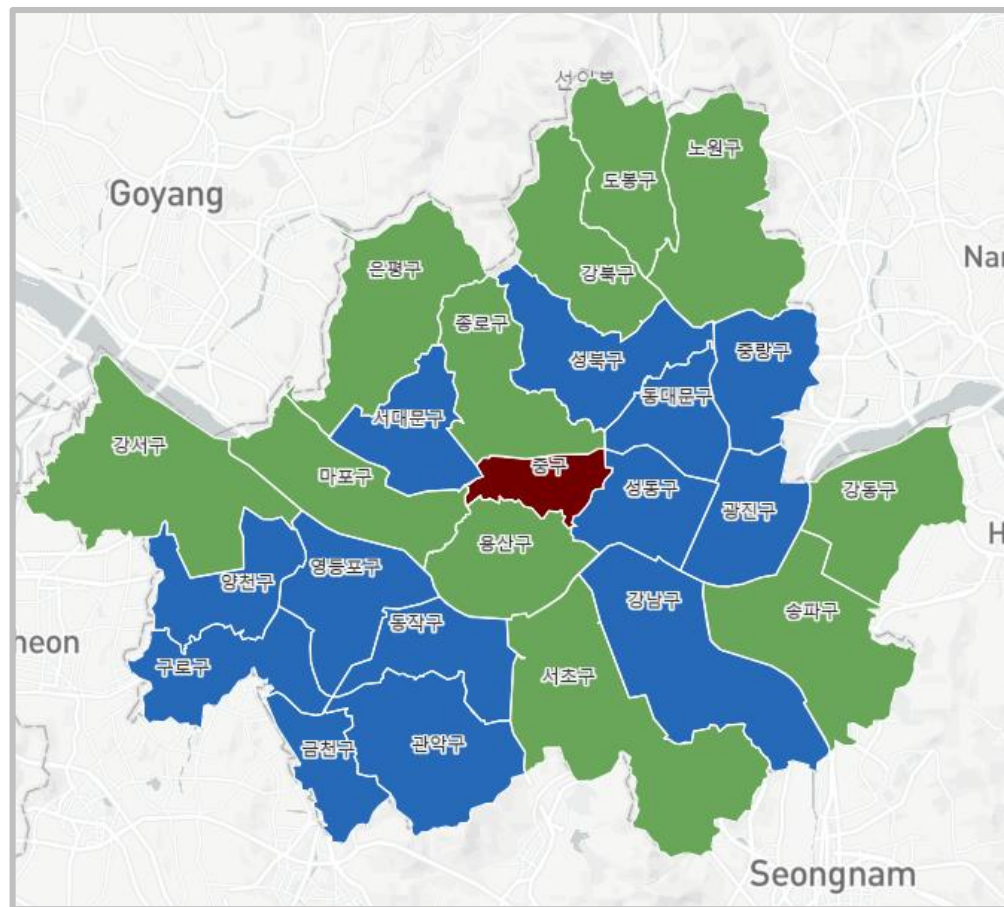


CCTV

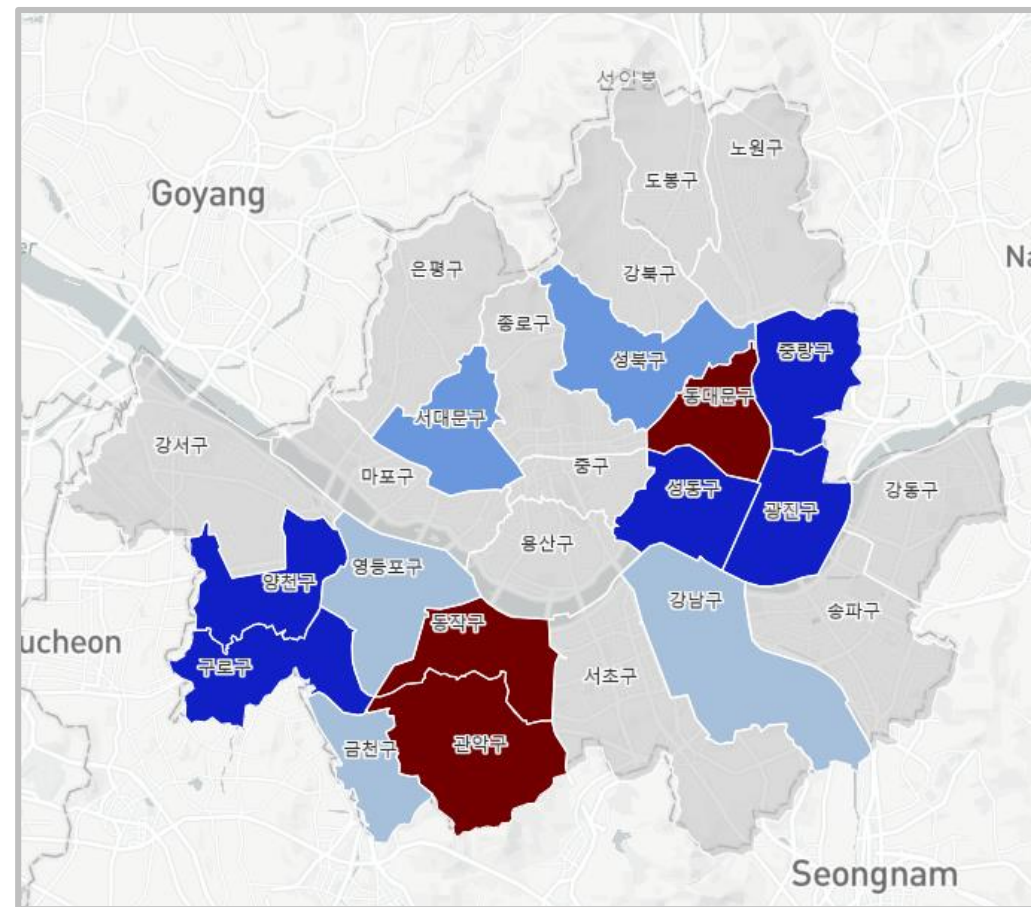
: cctv

분석 결과 군집화 결과 시각화

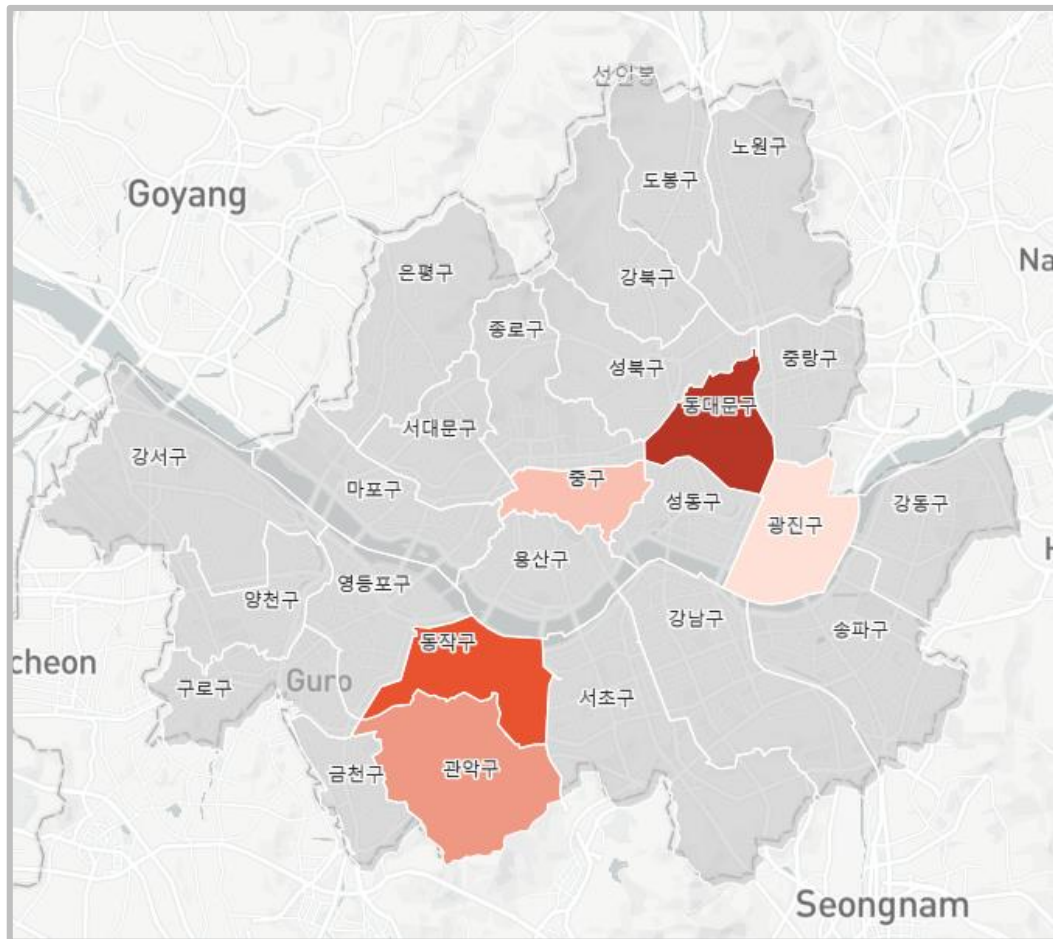
<1차 군집분석>



<2차 군집분석>



분석 결과 위험 지역 순위화



□ 지역안전지수 산출식

안전지수 = 100 - (위해지표 + 취약지표 - 경감지표)

$$= 100 - \left\{ \sum_{i=1}^n (\omega_i \times H_i) + \sum_{j=1}^m (\alpha_j \times C_j) - \sum_{k=1}^o (\beta_k \times M_k) \right\}$$

ω_i : 위해지표별 가중치

H_i : 위해(harm)지표 점수

α_j : 취약지표별 가중치

C_j : 취약(cause)지표 점수

β_k : 경감지표별 가중치

M_k : 경감(mitigation)지표 점수

rank	district
1	동대문구
2	동작구
3	관악구
4	중구
5	광진구

위험 지역 순위화 결과 값과

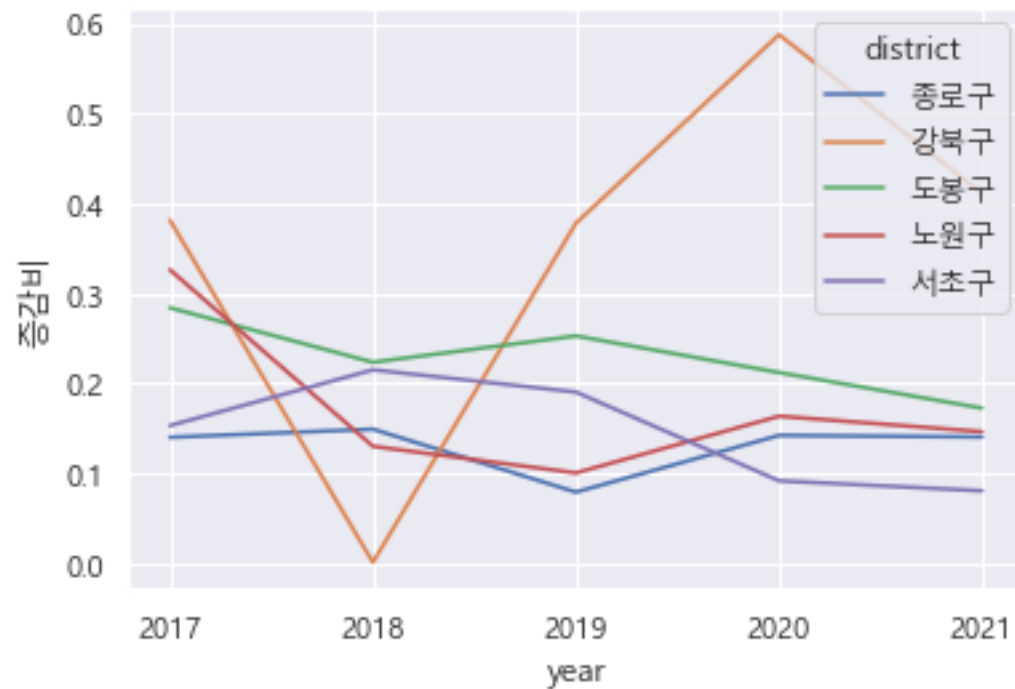
Kmeans 군집 분석을 통해 나온 결과 값 유사

+

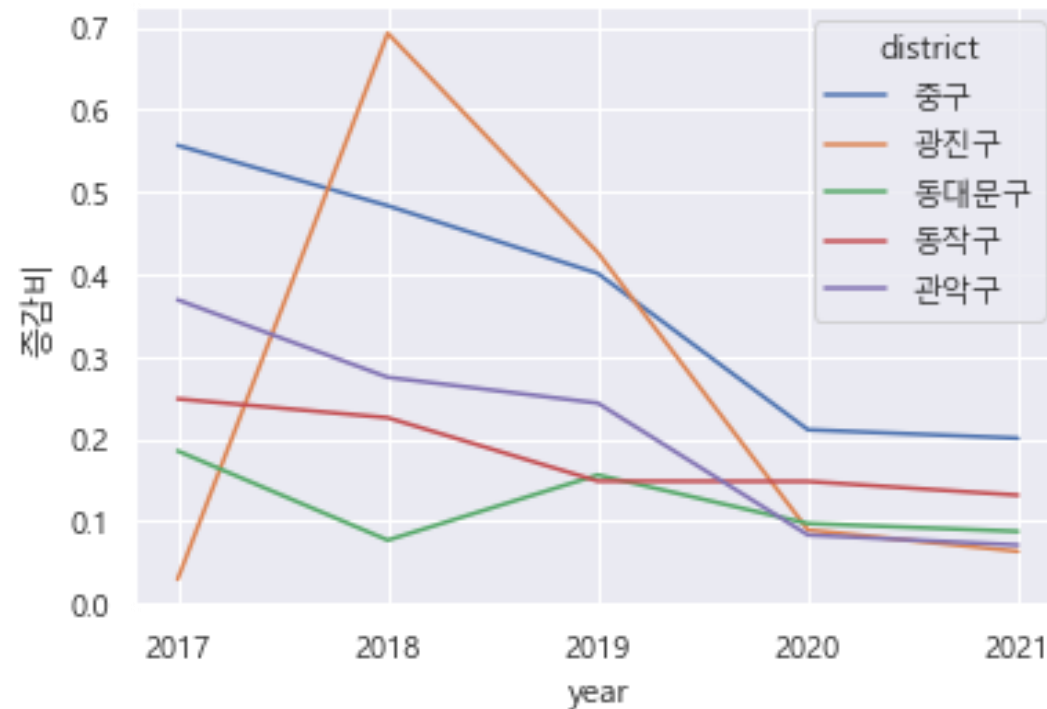
분석 결과

cctv 설치 필요 지역

<Safe_district>



<Danger_district>



Safe_district : 종로구, 강북구, 도봉구, 노원구, 서초구

Danger_district : 동작구, 동대문구, 관악구, 중구, 광진구

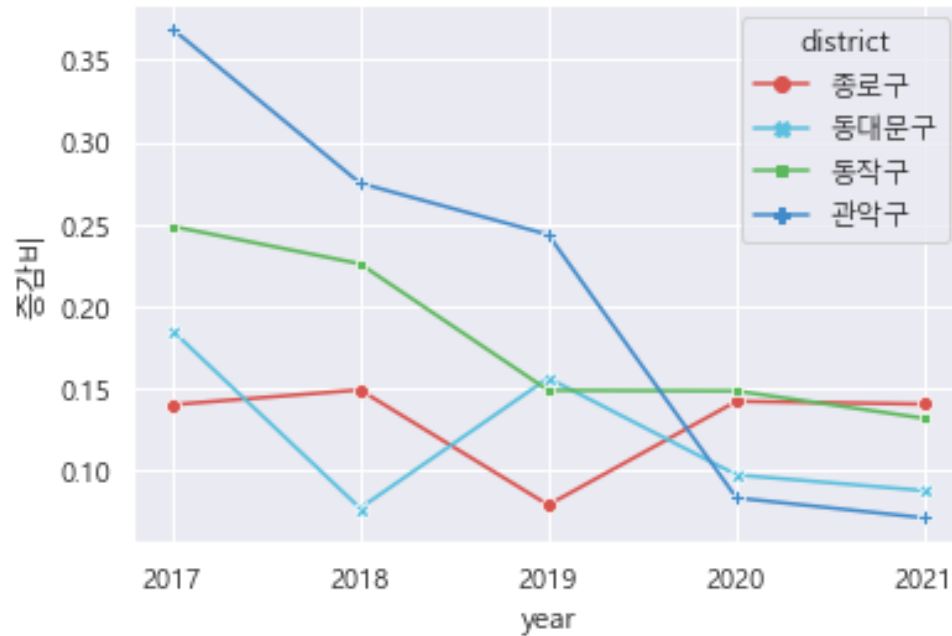
+

분석 결과

cctv 설치 필요 지역

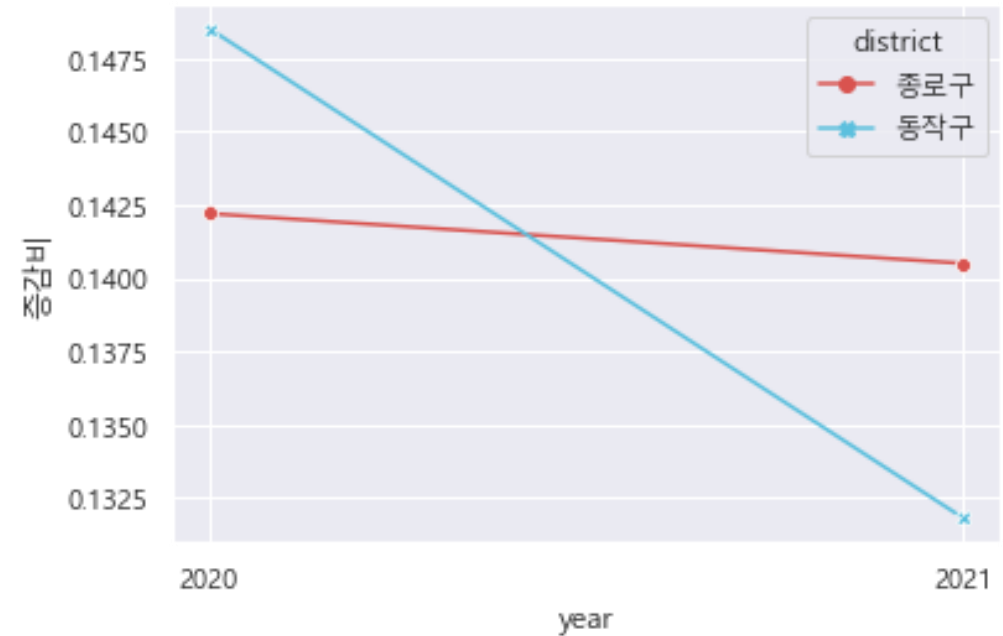
Safe_district : 종로구

Danger_district : 동작구, 동대문구, 관악구



Safe_district : 종로구

Danger_district : 동작구



위험 지역의 CCTV 증설이 점점 감소되고 있다.
반면 안전 지역의 CCTV 증설은 꾸준한 비율을 보이고 있다.

CCTV의 효율적인
자원분배를 위한 정책 필요

+

시사점 및 개선점

+

시사점

✓ 여성 1인 가구의 안전한 주거 환경을 위한 실효성 있는 행정 정책을 기대

-> 안전지역, 역으로는 위험지역을 순위로 제시함으로써 치안을 우선 강화할 지역을 선정할 경우 참고자료로 활용될 수 있음

-> 방범용 CCTV 증설 시 위험지역을 우선적으로 고려하는 정책 수립이 절실히 요구됨을 시사

✓ 여성 1인 가구들이 안전한 거주지를 선택하는데 참고 자료로 활용

-> 안전지역, 역으로는 위험지역을 순위로 제시함으로써 거주지 선택 시 참고 자료로 활용될 수 있음

✓ 해당 구 거주민들의 안전한 거주 환경 조성을 위한 참고 자료로 활용

-> 안전지역, 역으로는 위험지역을 순위로 제시함으로써 일반인들도 거주지 선택 시 참고 자료로 활용될 수 있음

+

개선점

- ✓ 유흥주점 같은 유의미한 변수에 경찰 출동 수를 늘리는 등의 정책 제안도 고려해 볼 수 있음
- ✓ 민간이 설치한 집 앞 또는 상가 내 감시카메라가 방범용 CCTV 역할을 하는 경우도 있어 공공데이터만 바탕으로 CCTV에 대한 정책을 제안하는 것이 한계가 있음
- ✓ 위험지역 선정에 있어서 추가적인 변수를 더 고려해보고 싶었음
- ✓ 이론적 지식을 강화하여 좀 더 체계적이고 데이터에 적합한 분석기법들을 적용할 필요성을 느낌
- ✓ Interactive data visualization 기법을 적용하여 결과를 좀 더 효과적으로 보여줄 필요성을 느낌

+

시사점 및 개선점

통계 분석

- 기존 5개의 변수 외 별도의 변수를 고려하지 못해 아쉬움
- 경찰서 및 치안센터가 유의한 변수가 아님
- 통계분석 방법의 적합도에 대한 아쉬움

분석 결과

- 안전지역 대비 위험지역 CCTV설치가 높기 때문에 본 프로젝트 결론이 다소 미흡
- 분석 결과물에 비해 해석의 아쉬움



감사합니다.

+

참조자료

<R kmeans 평균 군집 분석 >

<https://www.youtube.com/watch?v=uW9Zbhh1Dww>

<R 지도 그리기 참조>

<https://givitallugot.github.io/articles/2020-03/R-visualization-1-seoulmap>

<참조 논문>

범죄 분야 지역안전지수 산출방식에 관한 연구 - 이철희 (고려대학교 정책대학원 데이터정보학과 국내석사)

국가 및 지역 안전역량 진단 시뮬레이션 기법 개발 국립재난안전연구원

1인 가구와 범죄발생에 관한 연구: 서울시 25개 자치구 패널 자료를 중심으로 장진희(한양대학교 경제학 박사)

+