A young girl with dark hair is crying, her face partially obscured by smoke. In the background, two hands are visible, each holding a lit cigarette. The smoke is drifting towards the girl's face, creating a somber and distressing atmosphere. The overall image serves as a background for a public health poster.

국민의 건강을 위협하는 간접흡연 피해를 줄이자!

흡연 구역 최적 입지 선정

- 공공 빅데이터를 활용한 입지 선정 모델링 -

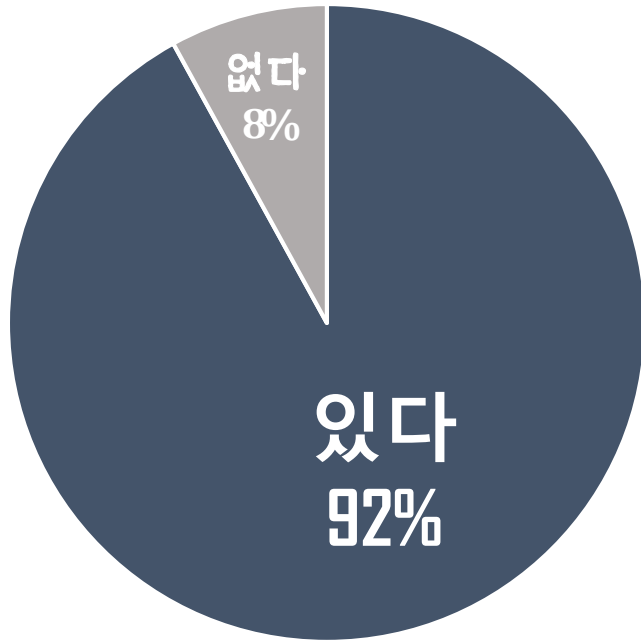
혁신멀캠팀

주영하
오영권

Contents

- 1** 상황분석
| 해결이 필요한 사회 문제 : '간접흡연'
- 2** 문제정의
| 간접흡연 원인 분석, 제안 배경
- 3** 분석 아이디어
| 문제 해결을 위한 분석 아이디어, 분석 순서도
- 4** 분석 과정
| 전처리 과정, 최적 입지 선정 모델링 과정
- 5** 기대효과
| 기대효과

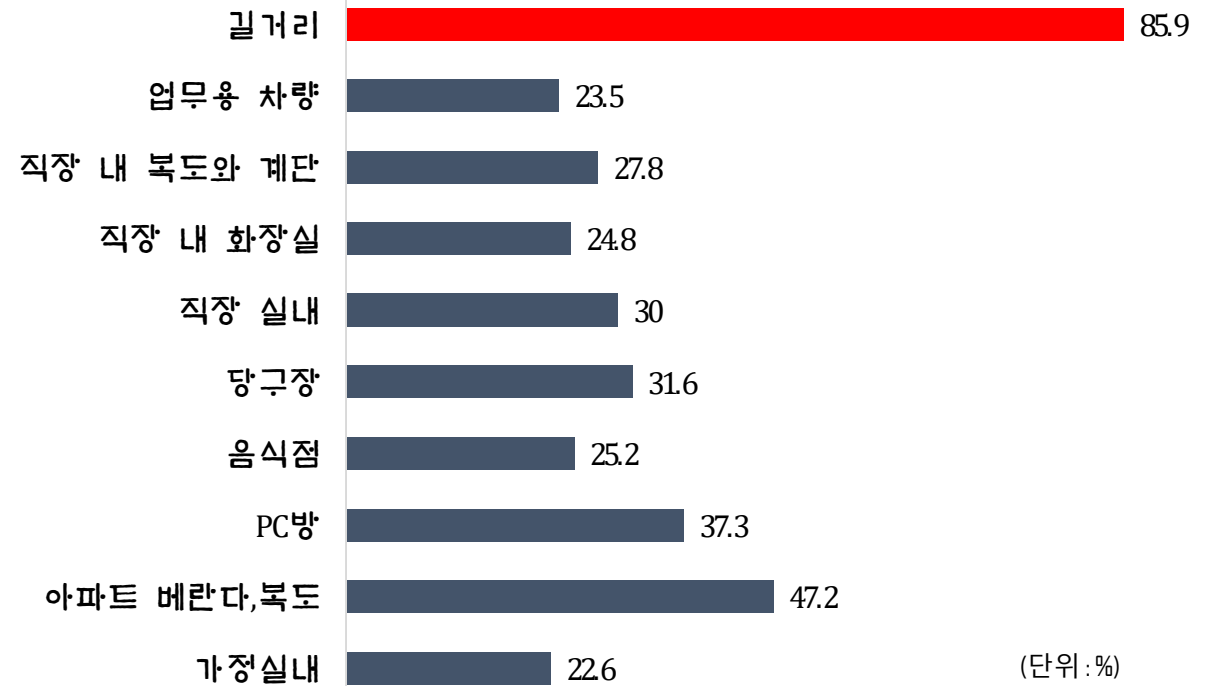
최근 한 달 사이 간접흡연 경험



Q: 최근 한 달 사이 본인이 원치 않지만 다른 사람이 피우는 담배연기를 맡은 적이 있습니까?

2019년 11월 전국 만 19세 이상 성인남녀 925명 대상 조사결과
자료: 한국리서치

한 달 기준 간접흡연 경험 장소



2018년 10~11월 국내 19~49살 성인 남녀 5280명 대상 간접흡연 피해장소 조사결과
자료: 한국보건사회연구원

많은 시민들이 간접 흡연에 노출되고 있으며, 그 중 길거리 간접 흡연이 가장 심각하다.

간접 흡연의 폐해

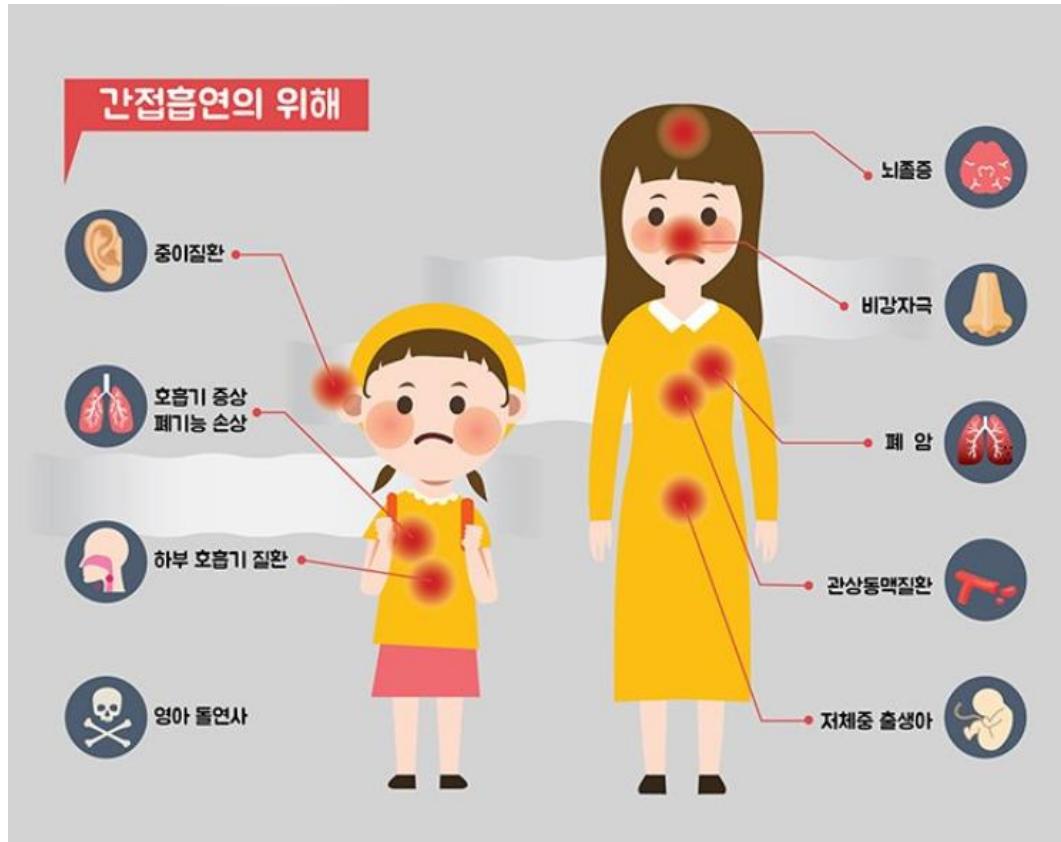
상황분석

문제정의

아이디어

알고리즘 및 모델
링

기대효과 및 한계
점



이미지 출처: 국가암 정보센터 NAVER 포스트

주류연 대비 부류연 독성 성분

일산화탄소	8배
암모니아	73배
디메틸나이트로소아민	52배
메틸나프탈렌	28배
아닐린	30배
나프탈아민	39배

*주류연: 흡연자가 빨아들이는 연기

*부류연: 담배가 타면서 나는 연기

출처: 매일경제 '가족, 동료 무슨 죄라고.. 간접흡연에 쓰러져간다.' (김기철, 이새봄 기자)

간접 흡연은 국민의 건강을 위협하고 있으며 해결해야 할 과제이다.

길거리 흡연 원인

상황분석

문제정의

아이디어

알고리즘 및 모델

기대효과 및 한계

링

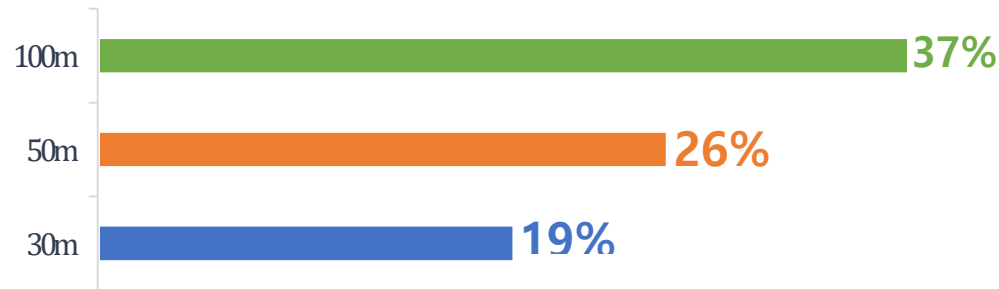
점

흡연 전 금연 구역 여부 확인



2019년 11월 전국 만 19세 이상 성인남녀 흡연자 245명 대상 조사결과
자료: 한국리서치

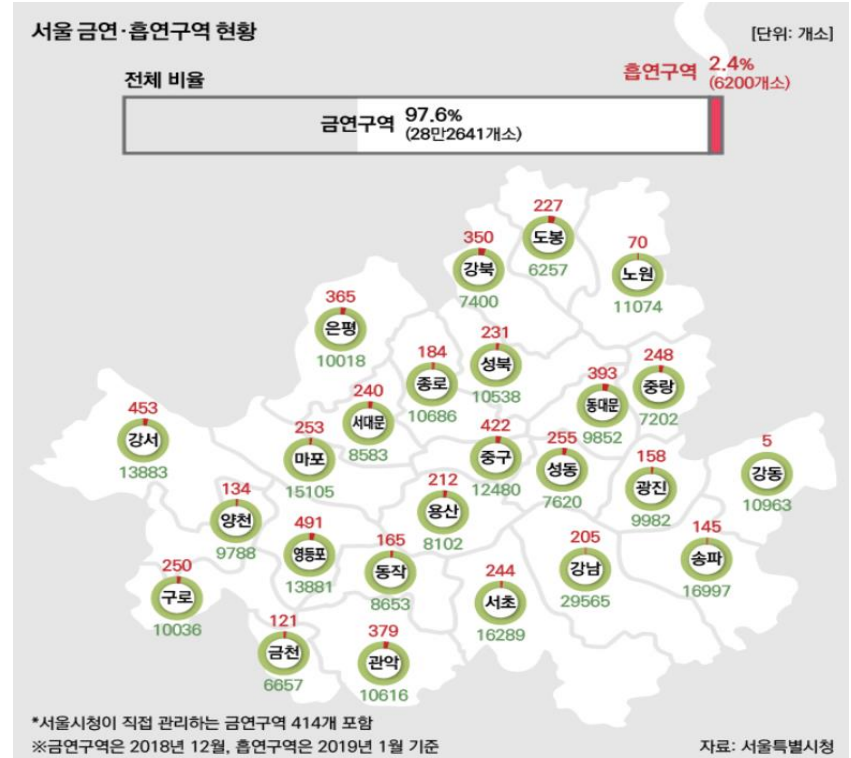
흡연을 위해 이동 가능한 거리



2017년 전국 성인 흡연자 103명 대상 조사결과
자료: 금연구역과 흡연공간의 분석을 통한 흡연공간과 서울시민의 행복증진의 인과관계 연구

흡연자의 대다수가 흡연 전 금연 구역을 확인하며
흡연 가능 구역만 있다면 다소 먼 거리도 이동하고자 한다.

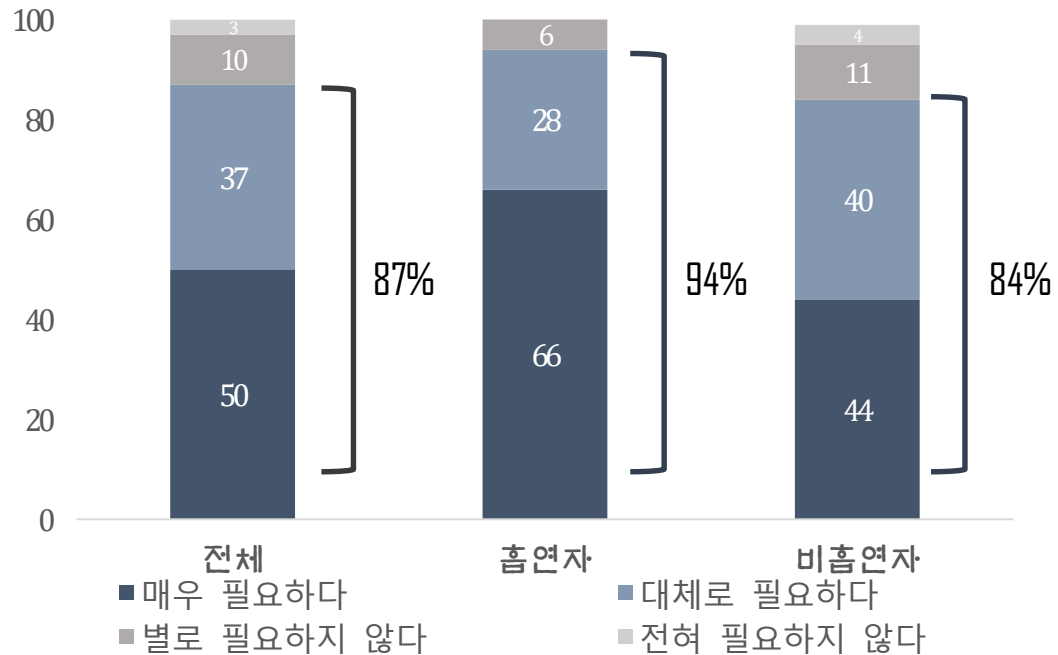
극도로 부족한 흡연 구역



서울시의 금연구역은 약 28만 개 이상
반면에, 흡연구역은 약 6200개로 금연구역 대비 2.4%에 불과하다.

무분별한 길거리 흡연의 원인은 흡연 공간의 부족이다.

흡연 구역 확대 동의 여부

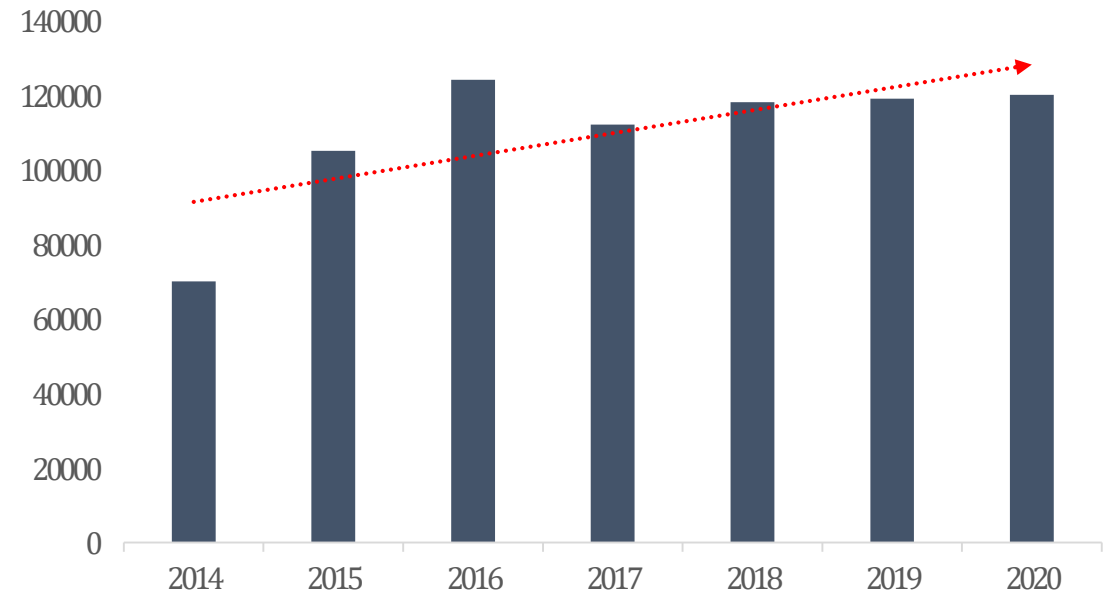


2019년 11월 전국 만 19세 이상 성인남녀 흡연자 245명 대상 조사결과
자료: 한국리서치

흡연자 뿐만 아니라 비흡연자도 흡연 구역 확대에 동의하는 비율이 높다.

담배 관련 제세 부담금

(단위: 억 원)



2019년 11월 전국 만 19세 이상 성인남녀 흡연자 245명 대상 조사결과
자료: 한국리서치

흡연자의 제세 부담금은 증가하는 추세이나 납부하는 세금에 비해 혜택이 없다는 불만이 꾸준히 제기되어 왔다.

공공 이익 실현을 위한 흡연 구역 확대가 필요하다.

DATA 공공데이터포털
.GO.KR



Raw Data

금연구역 수 인구 수

면적 흡연율 유동인구 ...

파생 변수

인구 밀도 총 매출 ...

공공 빅데이터

Algorithm

$$\begin{aligned} \text{Subject to } \text{Min } & \sum_i \sum_j h_i d_{ij} y_{ij} \\ & \sum_j y_{ij} = 1 \quad (\text{for all } i) \\ & \sum_j x_j = p \\ & y_{ij} \leq x_j \quad (\text{for all } i, j) \\ & y_{ij} \in \{0, 1\} \quad (\text{for all } i, j) \\ & x_j \in \{0, 1\} \quad (\text{for all } j) \end{aligned}$$

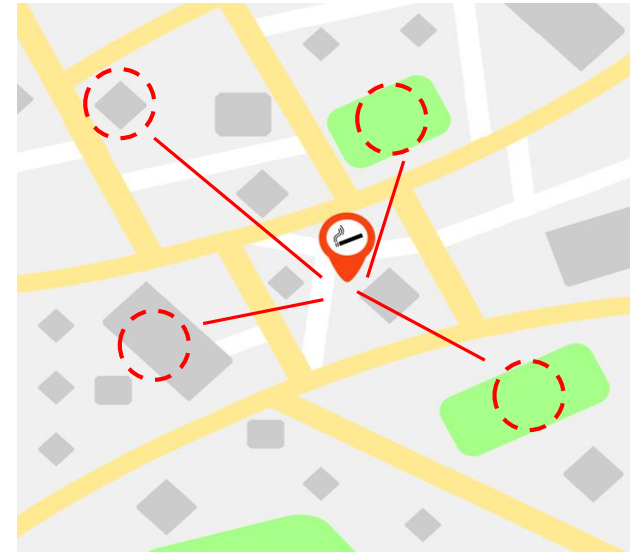
Tool



Google APIs

알고리즘 및 모델링

[구역 개수 최소화
최대 흡연 수요 커버]



최적 입지 선정

공공 빅데이터를 활용한 흡연 구역 최적 입지 선정 아이디어를 제안한다.

순서도

상황분석

문제정의

아이디어

알고리즘 및 모델링

기대효과 및 한계점

링

점

데이터 수집 및 전처리

각 자치구 별 데이터 수집



전처리 및 파생 변수 생성



1. 금연 구역 수
2. 흡연 구역 수
3. 인구 수
4. 인구 비율
5. 면적
6. 흡연률
7. 유동인구
8. 점포 수
9. 총 매출
10. 흡연 단속 건수
11. 인구 밀도
12. 흡연 구역 대비 금연구역
13. 흡연인구

우선 입지 자치구 선정

클러스터링 기법 3가지

K-means

Gaussian
Mixture

Hierarchical
Clustering



하이퍼 파라미터(K값) 별
Silhouette score 확인



Score 계수가 가장 높은
클러스터링 기법 채택



클러스터 별 특성 확인



Target Cluster(자치구)
선정



우선 적용 자치구 선정

최종 후보지 선정

보안등 위치를
흡연 구역 후보지로 설정



학교, 어린이보호구역 등
주요 시설 반경
300m를 설치 불가능
지역으로 설정



*Buffer*를 활용해
보안등(후보지) 필터링



Hierarchical Clustering



각 구 별 최종 후보지
300개 선정

입지 선정 모델링

LSCP

P-median 모델에
입력할 *P*값 도출



P-median

각 수요지에
*가중치*를 적용하여
최종 입지 선정



결과 시각화
&
최종 입지 선정

각 자치구별 유동인구를 지하철 승하차 인원과 버스 승하차 인원의 합으로 사용

CSV

서울시 지하철 호선별 역별 승하차 인원 정보

CSV

전국 도시철도 역사 정보 표준데이터



자치구 별 승하차 인원 그룹화



	gu_name	sub_pop
0	강남	807503
1	강동	217104
2	강서	206906
3	강북	125649
4	강원	262776

CSV

서울시 버스 노선별 정류장별 승하차 인원 정보

CSV

서울시 버스 정류소 좌표 데이터

- 구글 API 활용 좌표 데이터를 주소 데이터로 변환



자치구 별 승하차 인원 그룹화



	gu_name	bus_pop
0	강남	44120
1	강동	24997
2	강서	28720
3	강북	29781

최종 데이터

	gu_name	sub_pop	bus_pop	gu_pop
0	강남	807503	44120	851623
1	강동	217104	24997	242101
2	강서	206906	28720	235626

전처리 과정

변수 설정

Raw Data

공공 데이터 포털, 서울 열린 데이터 광장 등 Raw Data 수집

금연구역 수

인구 수

면적

흡연율

유동인구

점포 수

평균매출

...

파생 변수

Raw Data를 가공해 여러 파생 변수 생성

금연구역 대비 흡연구역

인구 밀도

총 매출

흡연 인구

금연구역 / 흡연구역

인구 수 / 면적

점포 수 x 평균매출

흡연율 x 인구 수

병합 결과

	자치구명	금연구역 수	흡연구역 수	인구 수	면적	흡연율	pop	점포수	합계	총	평균	매출	흡연단속건수	인구 밀도	흡연구역/금연구역	총
0	종로구	10686	184	145512	23.91	18.6	378475	3013	2146.222702	969	6085.821832	0.017219	6.466569			
1	중구	12480	422	122858	9.96	23.0	329859	2709	2516.067183	1786	12335.140562	0.033814	6.816026			
2	용산구	8102	212	224145	21.87	17.7	249756	2485	2267.656338	191	10248.971193	0.026166	5.635126			
3	성동구	7620	255	287820	16.86	19.5	453173	2318	2549.215272	950	17071.174377	0.033465	5.909081			

데이터 스케일링

	자치구명	금연구역 수	흡연구역 수	인구 수	면적	흡연율	pop	점포수	합계	총 평균 매출	흡연단속건수	인구 밀도	흡연구역/금연구역
0	종로구	10686	184	145512	23.91	18.6	378475	3013	2146.222702	969	6085.821832	0.0176	
1	중구	12480	422	122858	9.96	23.0	329859	2709	2516.067183	1786	12335.140562	0.0338	
2	용산구	8102	212	224145	21.87	17.7	249756	2485	2267.656338	191	10248.971193	0.0267	

각 변수 별 데이터 스케일 차이가 큼

스케일링(정규화) 작업 수행

Z-score Scaling

데이터와 평균의 차를 표준편차로 나누어 주는 Z-score scaling 수행
(이상치를 처리하기에 용이)

$$x'_i = \frac{x_i - m}{\sigma}$$

스케일링 결과

금연구역 수	흡연구역 수	인구 수	면적	흡연율	pop	점포수	합계	총 평균 매출	흡연단속건수	인구 밀도	흡연구역/금연구역
-0.128781	-0.543064	-1.892137	-0.032841	-0.097847	0.587908	0.096652	-0.286454	-0.240485	-2.304389		-0.5816
0.254308	1.476456	-2.073629	-1.562010	1.638141	0.248524	-0.163644	1.207911	-0.018648	-0.924337		0.8071
-0.680565	-0.305474	-1.262170	-0.256462	-0.452935	-0.310666	-0.355441	0.204201	-0.451733	-1.385031		0.1671

STEP 1

K-means, Hierarchical, Gaussian Mixture 클러스터링

STEP 2

Silhouette Score 측정, Silhouette Score가 가장 높은 클러스터링 방법 채택

STEP 3

클러스터 별 특성 확인 및 클러스터 선정

STEP 4

우선 적용 자치구 최종 선정 : RANK 함수 활용

우선 적용 자치구 선정

상황분석

문제정의

아이디어

알고리즘 및 모델

기대효과 및 한계

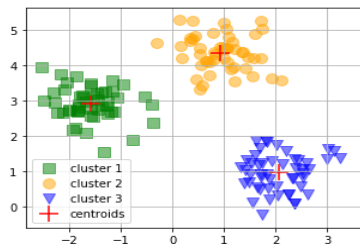
링

점

STEP 1

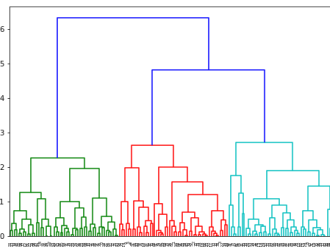
클러스터링 수행

K-means



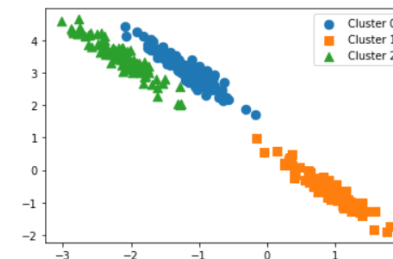
데이터 간의 거리가 **최소**가 되도록 하는
중심점을 생성하여 데이터들을 **K**개의
클러스터에 할당

Hierarchical



순차적, 계층적으로 유사한 그룹을
통합하여 **K**개의 클러스터에 할당

Gaussian Mixture



데이터가 **K**개의 정규분포로 이루어 졌다고
가정한 뒤 **조건부 확률**을 계산하여 제일 높은
확률 값을 가지는 클러스터에 데이터를 할당

STEP 2

Silhouette Score 측정

K	K-means Silhouette Score	Hierarchical Silhouette Score	Gaussian Mixture Silhouette Score
3	0.24556132419028065	0.1300329485640696	0.17411904471886128
4	0.11655300204889572	0.13455924352758977	0.20965242342762763
5	0.1398699105059083	0.14560055008750628	0.1645908785088056
6	0.17725357219124654	0.16959613586831357	0.20965242342762763
7	0.167388818284489	0.15097113684568864	0.18366501848533934
8	0.18541674217665455	0.1758135561520121	0.17231749830751966
9	0.19985801096669562	0.15971184089509763	0.12589503741233032

Silhouette Score가 가장 높은
K값이 3인 **K-means** 클러스터링 채택

우선 적용 자치구 선정

상황분석

문제정의

아이디어

알고리즘 및 모델

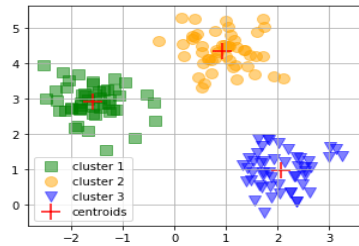
기대효과 및 한계

링

점

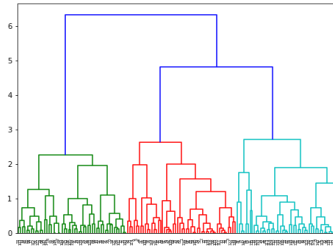
STEP 1 클러스터링 수행

K-means



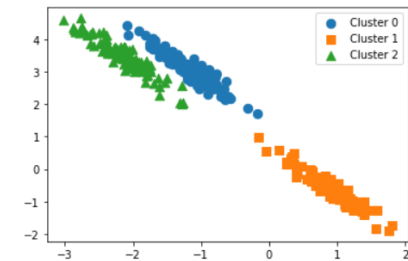
데이터 간의 거리가 **최소**가 되도록 하는
중심점을 생성하여 데이터들을 **K**개의
클러스터에 할당

Hierarchical



순차적, 계층적으로 유사한 그룹을
통합하여 **K**개의 클러스터에 할당

Gaussian Mixture



데이터가 **K**개의 정규분포로 이루어 졌다고
가정한 뒤 **조건부 확률**을 계산하여 제일 높은
확률 값을 가지는 클러스터에 데이터를 할당

STEP 2 Silhouette Score 측정

K	K-means Silhouette Score	Hierarchical Silhouette Score	Gaussian Mixture Silhouette Score
3	0.24556132419028065	0.1300329485640696	0.17411904471886128
4	0.11655300204889572	0.13455924352758977	0.20965242342762763
5	0.1398699105059083	0.14560055008750628	0.1645908785088056
6	0.17725357219124654	0.16959613586831357	0.20965242342762763
7	0.167388818284489	0.15097113684568864	0.18366501848533934
8	0.18541674217665455	0.1758135561520121	0.17231749830751966
9	0.19985801096669562	0.15971184089509763	0.12589503741233032

Silhouette Score가 가장 높은
K값이 3인 **K-means** 클러스터링 채택

STEP 3

클러스터별 특성 확인 및 클러스터 선택

K-means 클러스터링(K=3) 결과 및 군집 별 평균 값 확인 및 클러스터 선정

	금연구역 수	흡연구역 수	인구 수	면적	흡연율	유동인구	점포수	합계	총 평균 매출	흡연단속건수	인구 밀도	흡연구역/금연구역	총 매출	흡연인구
y_km														
0	9086.470588	254.941176	331335.117647	19.707059	19.923529	255520.529412	2328.764706	2185.227357	709.294118	17238.348006	0.028368	5.112269e+06	65668.673882	
1	29565.000000	205.000000	533359.000000	39.500000	18.000000	851623.000000	6877.000000	2840.394067	1224.000000	13502.759494	0.006934	1.953339e+07	96004.620000	
2	14027.428571	237.285714	482314.285714	32.960000	16.357143	308712.428571	3719.571429	2205.527119	4726.428571	15209.452852	0.016434	8.208704e+06	79093.586000	

흡연 단속 건수를 중요 변수로 판단, 높은 수치를 보이는 2번 클러스터 선정

2번 클러스터로 분류된 자치구

	자치구명	금연구역 수	흡연구역 수	인구 수	면적	흡연율	유동인구	점포수	합계	총 평균 매출	흡연단속건수	인구 밀도	흡연구역/금연구역	총 매출	흡연인구	y_km
10	노원구	11074	70	513053	35.44	17.6	286822	2233	2122.677564	2693	14476.664786	0.006321	4.739939e+06	90297.328	2	
13	마포구	15105	253	369903	23.85	16.8	320237	4928	1953.550325	366	15509.559748	0.016749	9.627096e+06	62143.704	2	
15	강서구	13883	453	575875	41.44	18.9	235626	3623	2274.156224	1294	13896.597490	0.032630	8.239268e+06	108840.375	2	
18	영등포구	13881	491	377536	24.55	15.7	349884	4306	2025.815838	6778	15378.248473	0.035372	8.723163e+06	59273.152	2	
21	서초구	16289	244	416285	46.98	15.0	337567	3796	2601.816122	18728	8860.898255	0.014979	9.876494e+06	62442.750	2	
23	송파구	16997	145	659385	33.87	14.9	388750	4273	2449.571261	2343	19468.113375	0.008531	1.046702e+07	98248.365	2	
24	강동구	10963	5	464163	24.59	15.6	242101	2878	2011.102502	883	18876.087841	0.000456	5.787953e+06	72409.428	2	

STEP 4

우선 적용 자치구 최종 선정

RANK 함수 활용 각 변수 별 Rank 지수 계산

2번 클러스터 중 Rank 지수 합이 가장 높은 4개 자치구를 우선 입지 행정 구로 선정

RANK 함수

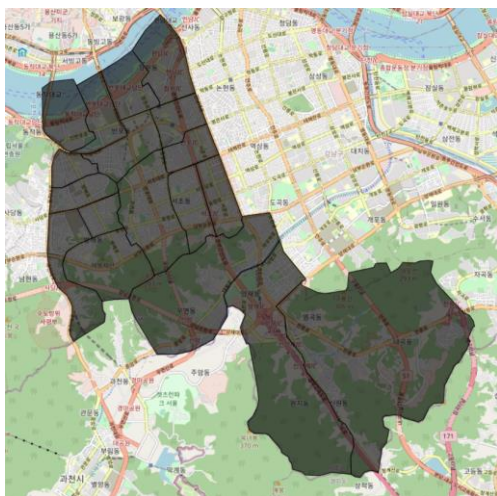
- 각 변수 별 수치를 비교하여 순위 점수 할당
 - 수치가 낮을수록 수요가 높을 것이라 판단되는 변수는 오름차순 처리
 - 총합 계산
- Ex) 흡연단속건수가 1위인 경우 7point, 2위인 경우 6point
흡연구역 수가 1위인 경우 1point 2위인 경우 2point

	자치구명	금연구역 수	흡연구역 수	인구 수	면적	흡연율	유동인구	점포수	합계	총 평균 매출	흡연단속건수	인구 밀도	흡연구역/금연구역	총 매출	흡연인구	y_km
10	노원구	11074	70	513053	35.44	17.6	286822	2233	2122.677564	2693	14476.664786	0.006321	4.739939e+06	90297.328	2	
13	마포구	15105	253	369903	23.85	16.8	320237	4928	1953.550325	366	15509.559748	0.016749	9.627096e+06	62143.704	2	
15	강서구	13883	453	575875	41.44	18.9	235626	3623	2274.156224	1294	13896.597490	0.032630	8.239268e+06	108840.375	2	
18	영등포구	13881	491	377536	24.55	15.7	349884	4306	2025.815838	6778	15378.248473	0.035372	8.723163e+06	59273.152	2	
21	서초구	16289	244	416285	46.98	15.0	337567	3796	2601.816122	18728	8860.898255	0.014979	9.876494e+06	62442.750	2	
23	송파구	16997	145	659385	33.87	14.9	388750	4273	2449.571261	2343	19468.113375	0.008531	1.046702e+07	98248.365	2	
24	강동구	10963	5	464163	24.59	15.6	242101	2878	2011.102502	883	18876.087841	0.000456	5.787953e+06	72409.428	2	

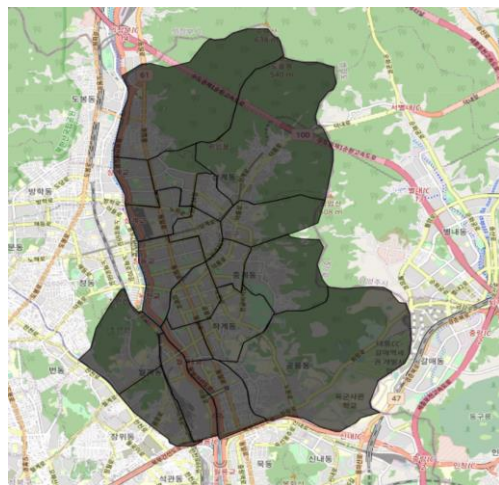
	자치구명	금연구역 수	흡연구역 수	인구 수	면적	흡연율	유동인구	점포수	합계	총 평균 매출	흡연단속건수	인구 밀도	흡연구역/금연구역	총 매출	흡연인구	y_km	rank_sum
10	노원구	11074	70	513053	35.44	17.6	286822	2233	2122.677564	2693	14476.664786	0.006321	4.739939e+06	90297.328	2	20.0	
13	마포구	15105	253	369903	23.85	16.8	320237	4928	1953.550325	366	15509.559748	0.016749	9.627096e+06	62143.704	2	17.0	
15	강서구	13883	453	575875	41.44	18.9	235626	3623	2274.156224	1294	13896.597490	0.032630	8.239268e+06	108840.375	2	16.0	
18	영등포구	13881	491	377536	24.55	15.7	349884	4306	2025.815838	6778	15378.248473	0.035372	8.723163e+06	59273.152	2	20.0	
21	서초구	16289	244	416285	46.98	15.0	337567	3796	2601.816122	18728	8860.898255	0.014979	9.876494e+06	62442.750	2	23.0	
23	송파구	16997	145	659385	33.87	14.9	388750	4273	2449.571261	2343	19468.113375	0.008531	1.046702e+07	98248.365	2	27.0	
24	강동구	10963	5	464163	24.59	15.6	242101	2878	2011.102502	883	18876.087841	0.000456	5.787953e+06	72409.428	2	17.0	

우선 적용 자치구 선정 결과

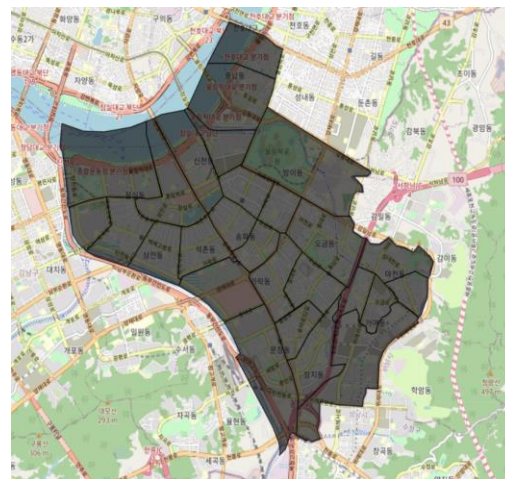
시각화를 위해 자치구 경계 데이터 결합



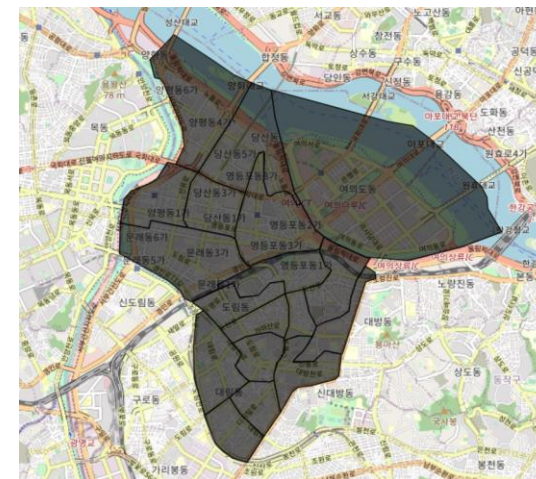
서초구



노원구



송파구



영등포구

STEP 1

데이터 구성 : 모델에 필요한 입지 후보지, 수요지, 가중치 데이터 구성 및 전처리

STEP 2

클러스터링을 통한 최종 후보지 선정

STEP 3

최종 입지 선정 : LSCP 모델, P-median 모델

STEP 4

입지 선정 결과 시각화 및 저장

STEP1 데이터 구성 : 입지 후보지

입지 후보지 : 보안등



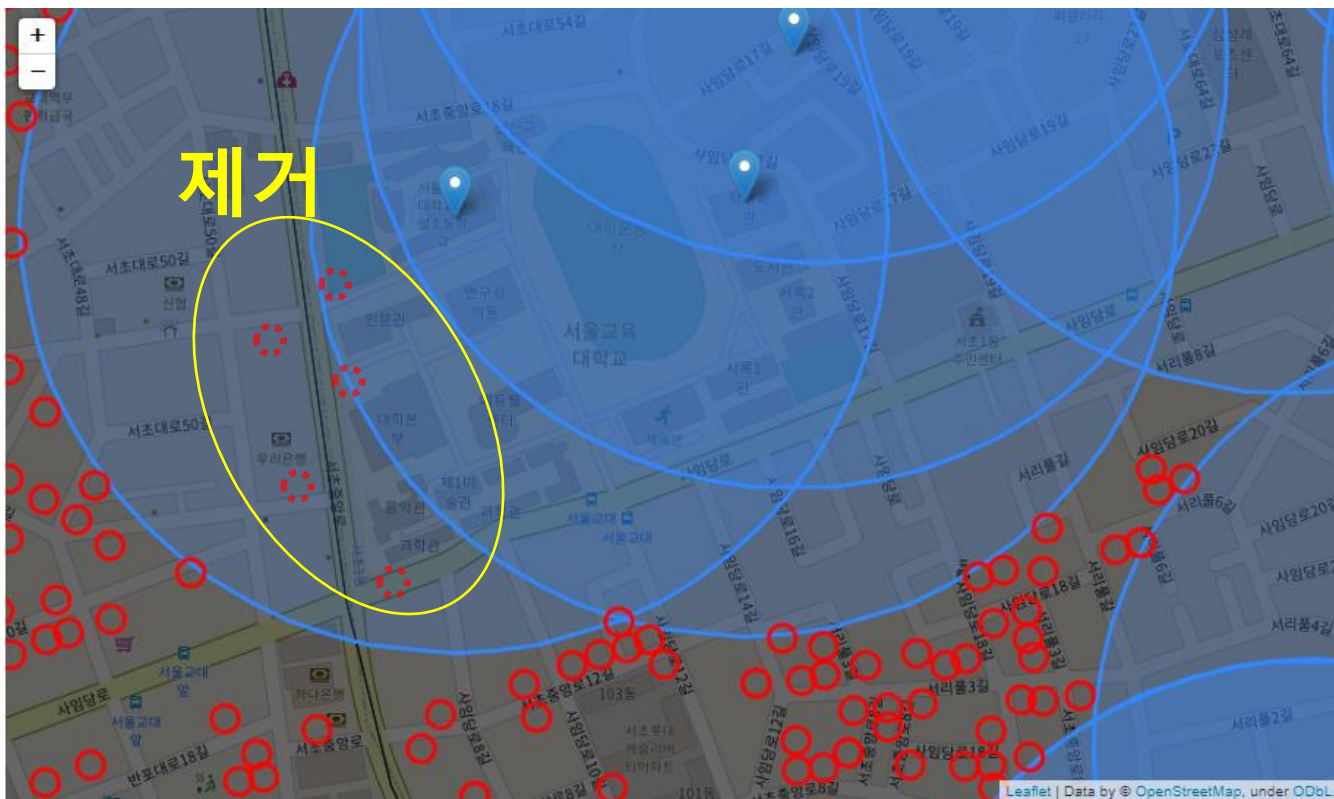
- 모든 상가 지역을 커버할 수 있으며 위치 데이터를 얻기 용이
- 공간의 크기와 특성 상 흡연 구역을 설치하기에 적합한 공간이라 판단
- 대로변이 아닌 골목 쪽에 주로 위치하는 보안등을 입지로 선정함으로써 보행자의 간접 흡연을 최소화

STEP1 데이터 구성: 입지 후보지 데이터 전처리

Buffer를 통해 흡연 구역 설치 불가 지역 필터링

학교, 어린이 보호구역 반경 300m, 기존 흡연 구역 반경 100m를 buffer로 설정하여 buffer 내에 있는 보안등을 후보지에서 제거

- Buffer: 데이터 분석에서 버퍼는 특정 개체로부터 일정 거리 안의 주변을 의미



학교 및 어린이 보호구역
주변 300m



기존에 설치된 흡연구
역 주변 100m



보안등

STEP 1

데이터 구성 : 수요지, 가중치

수요지 : 상가



- 조사 결과에 따르면 근무, 식사, 음주 시 흡연 욕구 발생이 가장 많았다.
- 회사 주변은 지정된 흡연 구역이 있기 때문에 수요지에서 회사 제외
- 식사, 음주 등의 활동은 일반적으로 상가에서 발생
→ 상가를 수요지로 설정

수요지 별 가중치

- 각 대분류 별 상가 수를 Rank화 하여 **가중치(Weights)**로 사용

- 카테고리 별 상가 수는 해당 카테고리를 방문하는 유동인구와 비례한다

- 따라서 카테고리 별 상가 수가 많으면 흡연 수요도 높을 것이라 판단.

데이터 구성: 수요지(상가) 데이터 전처리

CSV

소상공인시장진흥공단 상가(상권)정보

시설 내에 흡연부스, 흡연구역이 존재하거나, 사유지로 흡연을 허용하는 업종은 흡연 구역 수요지에서 제외

[x] 'PC/오락/당구/볼링 등', '경마/경륜/성인오락', '무도/유흥/가무', '연극/영화/극장', '호텔/콘도'

흡연 구역의 수요지로 볼 수 없는 시설 제외

[x] '학원-보습교습입시', '유아교육', '학원-어학'



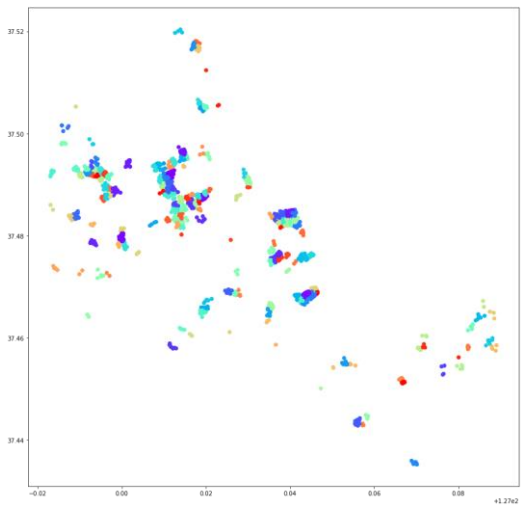
최종 수요지 데이터

[illegible]

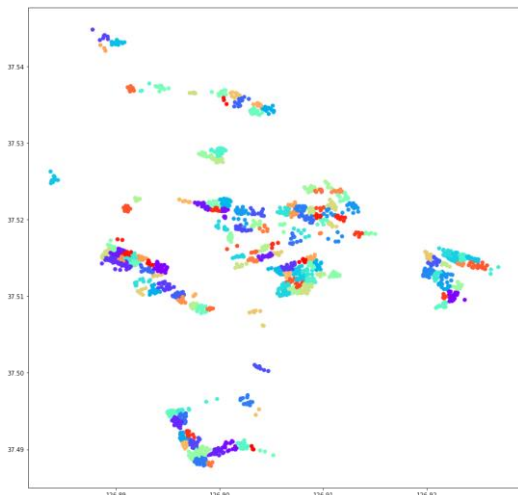
STEP 2 Hierarchical 클러스터링을 통한 최종 후보지 선정

- 1) 클러스터 수를 하이퍼파라미터 값으로 설정할 수 있는 Hierarchical 클러스터링 사용
- 2) K개의 클러스터를 추출하여 각 클러스터 중심점에서 가장 가까운 보안등을 최종 후보지로 선정.
- 3) 최적 클러스터 수 : 각 자치구 후보지(보안등) 100m 내에 평균 n개의 상가가 존재할 때, 각 후보지가 n개의 상가를 커버하도록 하는 후보지의 수를 최적 클러스터 수(K)로 선정

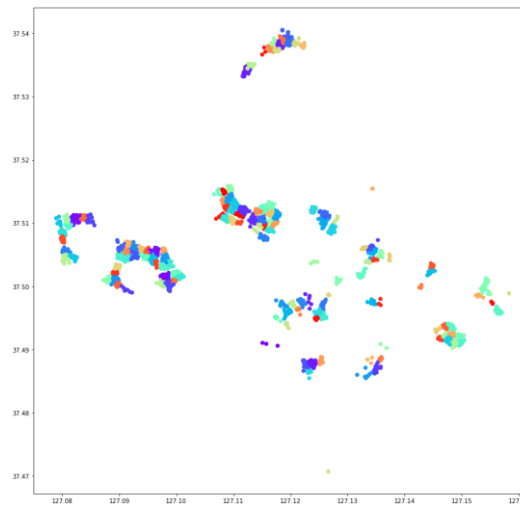
각 자치구 별 전체 후보지 100m 내 상가 수 / n = 클러스터 수(K)



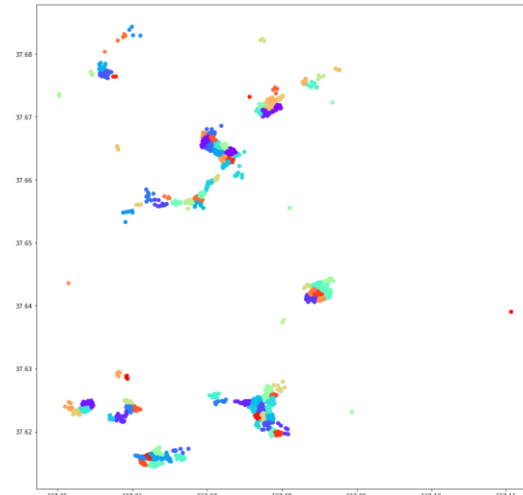
서초구 K : 168



영등포구 K : 156



송파구 K : 141



노원구 K : 105

STEP 3 최종 입지 선정 모델링

LSCP 모델

$$\begin{aligned} \text{Min} Z &= \sum_{j=1}^n X_j \\ \text{s.t.} \quad &\sum_{j=1}^n a_{ij} X_j \geq 1 \quad (\text{for all } i) \\ &X_j \in \{0, 1\} \quad (\text{for all } i) \end{aligned}$$

특정한 지리적 범위 내에 모든 수요가 커버 될 수 있도록 시설물의 수를 최소화하는 것을 목적으로 하는 모델

Input

입지 후보지, 수요지

Output

최적입지, 최소 시설 수

LSCP 모델의 최소 시설 수를 p-median의 p값으로 선정

P-median 모델

$$\begin{aligned} \text{Subject to Min} \quad &\sum_i \sum_j h_i d_{ij} y_{ij} \\ &\sum_j y_{ij} = 1 \quad (\text{for all } i) \\ &\sum_j x_j = p \\ &y_{ij} \leq x_j \quad (\text{for all } i, j) \\ &y_{ij} \in \{0, 1\} \quad (\text{for all } i, j) \\ &x_j \in \{0, 1\} \quad (\text{for all } j) \end{aligned}$$

후보입지가 주어져 있고 각 소비자(수요 발생지)에게 비용과 비용 거리가 주어져 있다고 가정할 때, 최소의 비용으로 모든 소비자의 수요를 충족시킬 수 있는 p개의 시설 설치 입지를 결정하는 모델

Input

입지 후보지, 최종 입지 수(p값), 수요지, 가중치

Output

최종 입지

STEP 4

모델링 결과 및 시각화

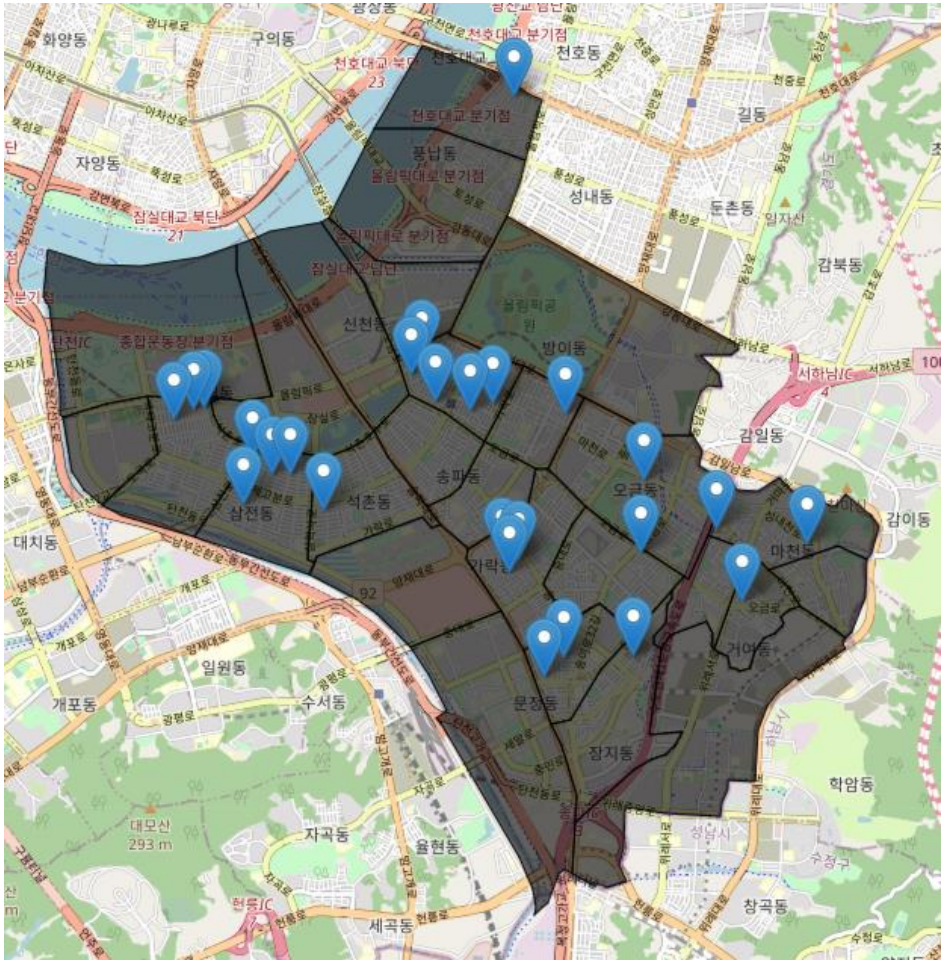


최종 입지



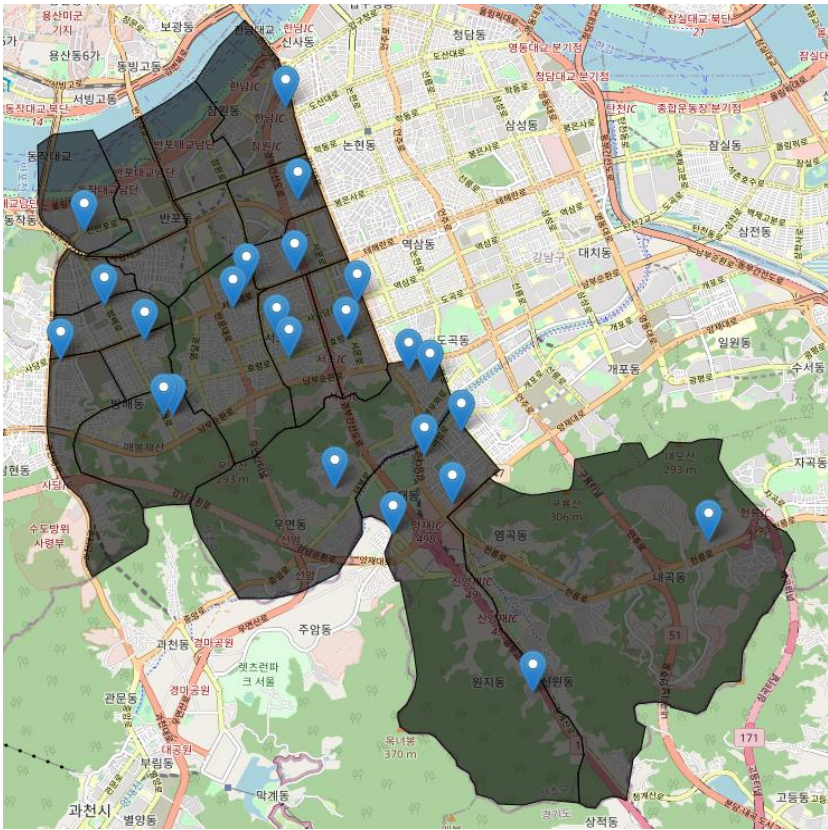
수요지(상가)

송파구 모델링 결과

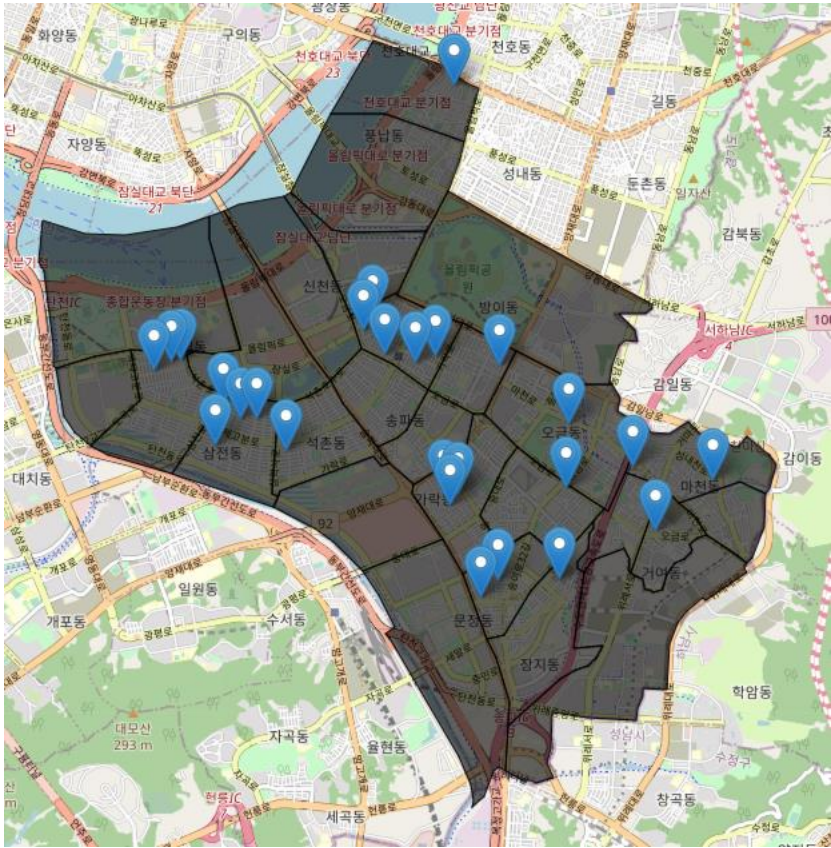


최종 입지 시각화

최종 입지 선정 결과

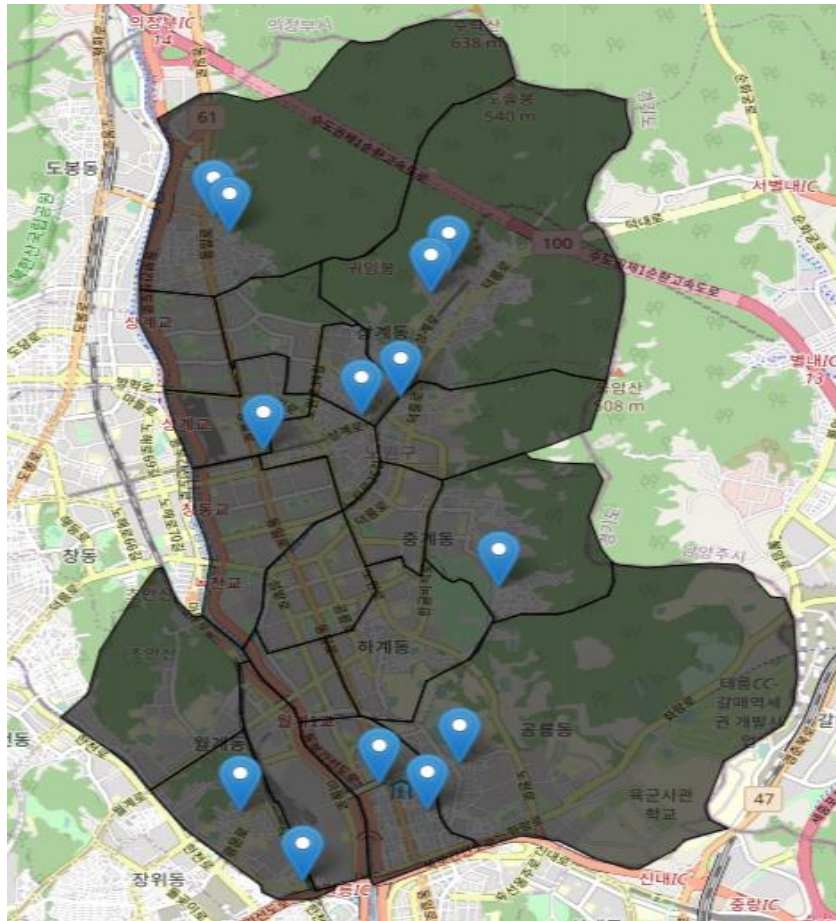


서초구

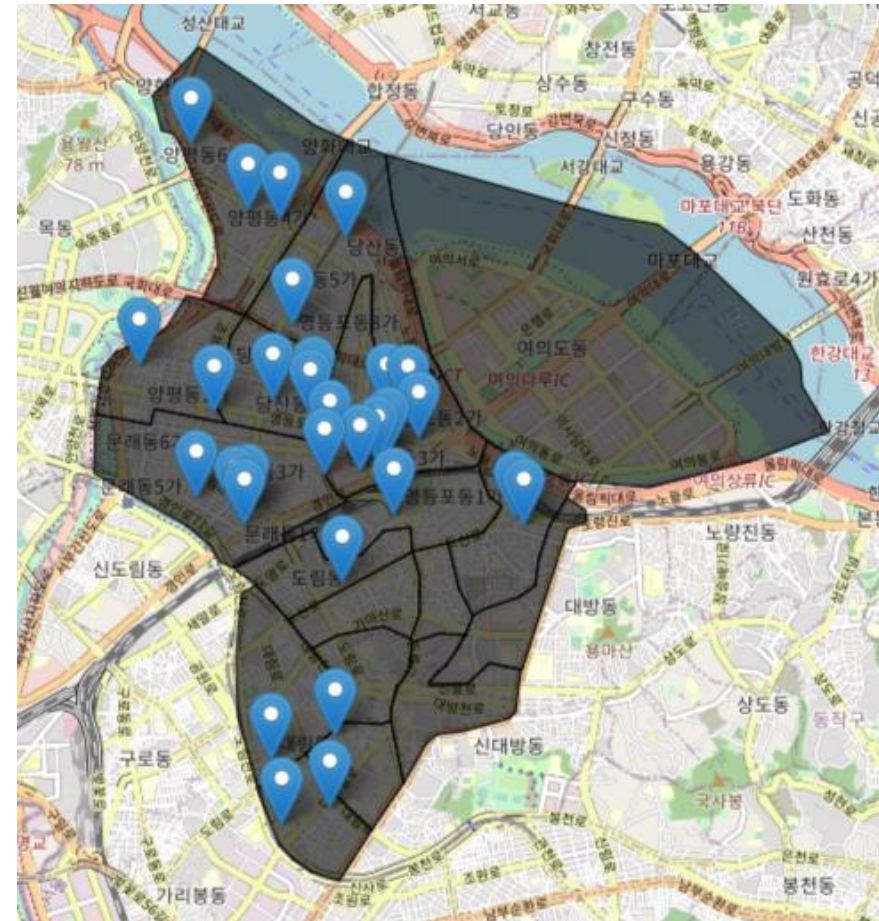


송파구

최종 입지 선정 결과



노원구



영등포구

1

최적 위치의 흡연 구역 선정을 통한 공공 이익 실현

- 적절한 위치의 흡연 구역으로 흡연자 복지 실현
- 비흡연자의 **간접 흡연** 피해 완화

2

흡연 구역 설치의 **딜레마** 완화

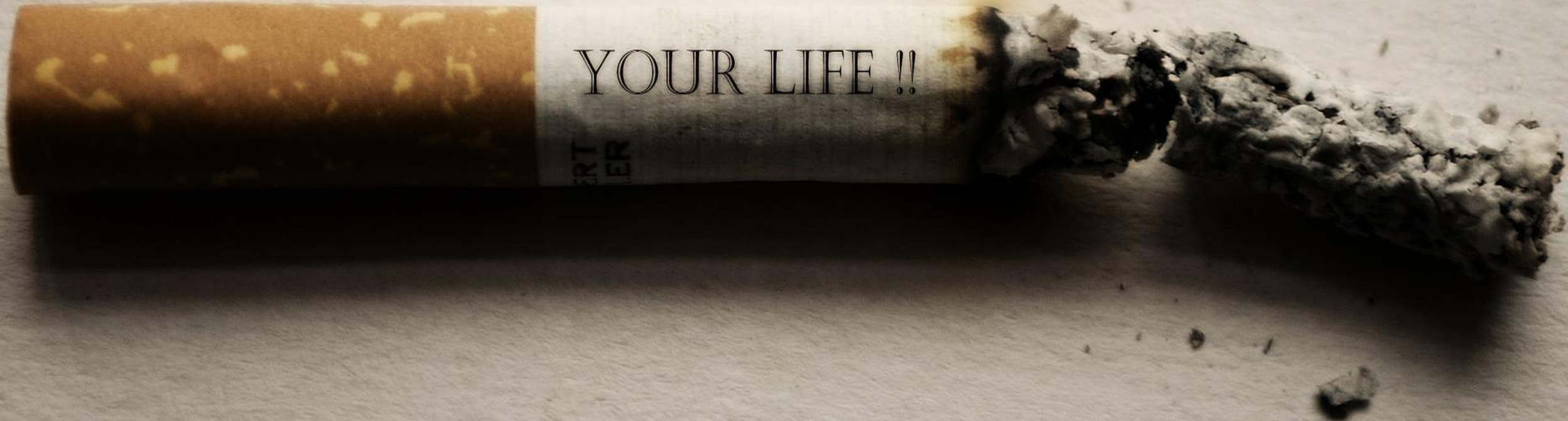
- '흡연 구역 설치 시 흡연자 건강 방치, 미 설치 시 간접흡연 피해 방치'라는 딜레마 존재
- **흡연 수요를 충족하는 최소 구역만 설치**하여 이를 완화

3

흡연 구역 설치의 최소화는 **운영 비용 절감**으로 연결

"가장 강한 금연정책은 명확한 흡연 구역 지정이다."

- 싱가포르 암 협회 -



감 사 합 니 다

전 국민이 간접 흡연에서 안전해지는 날이 오길 바랍니다!

한계점 및 보완가능성

상황분석

문제정의

아이디어

알고리즘 및 모델
링

기대효과 및 한계
점

한계점

서울 외 지역은 관련 데이터가 부족하여 서울에 한정된 입지 선정

보완 가능성

유사 데이터 확보가 가능하다면 전국으로 확장 가능

한계점

클러스터링 주요 변수 선정 시 일정 부분 주관적 판단

보완 가능성

충분한 데이터 셋이 추가 된다면 다중선형회귀 분석을 통해 주요 변수를 객관적으로 선정 가능

한계점

공개된 데이터들의 기준이 자치구별로 달라 데이터 간 차이 존재

보완 가능성

참고 문헌

- [1] 박동찬 외 2인, 금연구역과 흡연공간의 분석을 통한 흡연공간과 서울시민의 행복증진의 인과관계 연구, 2017 서울연구논문 우수작
- [2] 김민지, 보행 중 흡연 금지 억압인가? 예방인가?, 한화생명 블로그 Life & Talk, 2019.03.06
- [3] 박강서 외 2인, '흡연의 자유와 혐연의 권리 사이' - 간접흡연에 대한 인식 조사, 2019.11.14
- [4] 원승일 기자, 담뱃값 인상 논란 진행형... 지난해 담배 판매량 4% 늘었다, 2021.01.29
- [5] 김양중 기자, 간접흡연 피해 경험한 곳은?...86% "길거리", 2019.08.14

사용 데이터

데이터명	활용목적	출처
전국 도시철도 역사 정보 표준데이터	유동인구 데이터 구축	서울 열린 데이터 광장
서울시 지하철 호선별 역별 승하차 인원 정보	유동인구 데이터 구축	서울 열린 데이터 광장
서울시 버스 노선별 정류장별 승하차 인원 정보	유동인구 데이터 구축	서울 열린 데이터 광장
서울시 버스 정류소 좌표 데이터	유동인구 데이터 구축	서울 열린 데이터 광장
보안등 표준 데이터	후보지 설정	공공데이터포털
소상공인진흥공단_상가(상권)정보_서울	수요지 설정	소상공인진흥공단
서울특별시_자치구_어린이보호구역	Buffer 설정	공공데이터포털
한국교원대학교_초중고등학교위치	Buffer 설정	공공데이터포털
서울시 흡연단속건수 (2018)	변수 및 파생변수 구축	서울 열린 데이터 광장
면적(구)	변수 및 파생변수 구축	서울 열린 데이터 광장
구별 흡연율	변수 및 파생변수 구축	서울 열린 데이터 광장
서울특별시_행정구_금연구역/흡연구역	변수 및 파생변수 구축	공공데이터포털

분석 도구

