



서울시 주택 밀집 지역의 거점형 분리 배출 시설 입지 제안

8팀 마주연 신보람 신해빈 정유정



목차



01 분석 배경

02 데이터 소개

03 분석 방법

04 결론

01 분석 배경

01 분석 배경

■ 서울시 주택 밀집 지역 쓰레기 현황



서울시 아파트를 제외한 주택은 생활폐기물 배출 장소가 정해져 있지 않음

4



심각한 쓰레기 무단 배출 초래

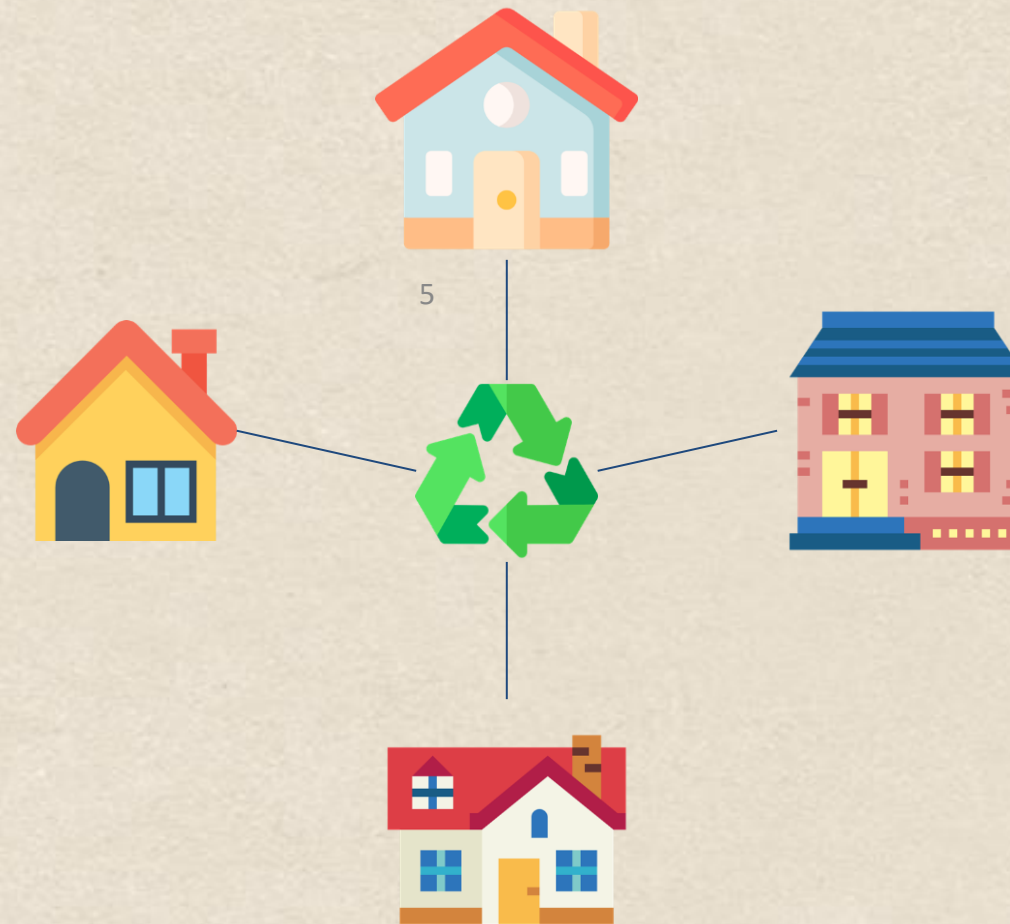


출처: "분리수거요? 양심을 지키고 싶어도 지킬 수 없는 환경이 문제죠" (skyedaily.com) 기사 사진

01 분석 배경

■ 분석 주제 : 서울시 주택 밀집 지역의 거점형 분리 배출 시설 입지 제안

주택의 위치를 고려하여 거점형 분리 배출 시설을 설치하면 쓰레기 무단 배출량이 감소하지 않을까?



01 분석 배경

■ 해외 사례 : 체코

코로나19 이후 쓰레기 배출량이 증가했지만 “분리수거 컨테이너”를 설치한 후 분리수거율 73%로 전년 대비 증가하였다.



2020년 기준 55만 8000개



주택 당 도보 2분 거리 (90m)



평균 112명 당 하나의 시설

높은 접근성

02 데이터 소개

02 데이터 소개

■ 생활폐기물 발생량

- 서울 열린데이터 광장에서 제공하는 「서울시 생활폐기물 발생량 및 처리현황 통계」
- 2020년 서울시 자치구별 생활폐기물 발생량 데이터 이용

통계



환경

활용사례(갤러리) 등록

URL 복사

목록 이동

서울시 생활폐기물 발생량 및 처리현황 통계

○ 통계개요

- * 통계명 : 생활폐기물 발생량 및 처리현황
- * 통계종류 : 서울시 생활폐기물의 발생량 및 처리현황을 제공하는 지정 · 조사통계
- * 근거법령 : 「폐기물관리법」에 의한 생활폐기물 발생 및 처리

출처: <https://data.seoul.go.kr/dataList/370/C/2/datasetView.do>

02 데이터 소개

■ 주택 수

- 서울 열린데이터 광장에서 제공하는 「서울시 주택종류별 주택(구별) 통계」
- 2020년 서울시 자치구별 단독주택+연립주택+다세대주택 수 데이터 이용

통계


주택/건설

활용사례(갤러리) 등록

URL 복사

목록 이동

서울시 주택종류별 주택(구별) 통계

○ 통계개요

- * 통계명 : 주택종류별 주택(구별)
- * 통계종류 : 서울시 주택현황을 주택종류별 · 자치구별로 제공하는 지정 · 조사통계
- * 근거법령 : 「인구주택총조사 규칙」(기획재정부령 제173호)

출처: <https://data.seoul.go.kr/dataList/11048/S/2/datasetView.do>

02 데이터 소개

■ 1인가구 평균 배달일수

- 서울 열린데이터 광장에서 제공하는 「서울 시민생활 데이터」

서울 시민생활 데이터란?

서울시와 SK텔레콤이 공공빅데이터와 통신데이터 가명결합을 통해 추정한 서울 행정동단위 성, 연령별 1인가구와 서울시민의 생활특성 정보

- 2022년 서울시 자치구별 월평균 배달 사용일수¹⁰ 데이터 이용

월평균 배달 사용일수 = 최근 3개월 총 배달 서비스 사용일수 / 3개월
(해당 날짜의 사용여부는 서비스의 접속 여부로 판단)


출처: <https://data.seoul.go.kr/dataVisual/seoul/seoulLiving.do>

02 데이터 소개

1인가구수

- 서울 열린데이터 광장에서 제공하는 「서울시 가구원수별 가구수(구별) 통계」
- 2020년 서울시 자치구별 1인가구수 데이터 이용

통계



복지

활용사례(갤러리) 등록

URL 복사

목록 이동

서울시 가구원수별 가구수(구별) 통계

○ 통계개요

- * 통계명 : 가구원수별 가구-읍면동(연도 끝자리 0,5), 시군구(그 외 연도)
- * 통계종류 : 조사, 지정통계
- * 근거법령 : 인구주택총조사 규칙


출처: <https://data.seoul.go.kr/dataList/10996/S/2/datasetView.do>

02 데이터 소개

■ 주민등록인구수

- 서울 열린데이터 광장에서 제공하는 「서울시 주민등록인구(월별_구별) 통계」
- 2020년 서울시 자치구별 주민등록인구수 데이터 이용
- 통계청의 「통계정보보고서」에서 '가구수 데이터가 2020년 11월 1일 기준'이라는 내용을 바탕으로 주민등록인구수 데이터도 2020년 10월 자료 이용

통계



일반행정

12

활용사례(갤러리) 등록

URL 복사

목록 이동

서울시 주민등록인구(월별_구별) 통계

서울시 통계정보시스템에서 제공하는 주민등록인구(월별_구별)에 대한 통계정보입니다. 서울시 주민등록인구(월별_구별)을 자치구별로 제공하는 일반
· 보고통계

02 데이터 소개

1인가구 비율

- 1인가구수 데이터와 주민등록인구수 데이터를 이용하여 1인가구 비율 계산
- 1인가구 비율 = 1인가구수 / 거주인구수

지역	1인가구수	거주인구수	1인가구 비율
종로구	25,983 ¹³	159,424	0.162980
중구	22,818	135,053	0.168956
용산구	36,881	245,009	0.150529
성동구	42,585	30,2134	0.140947
광진구	62,301	361,551	0.172316

02 데이터 소개

■ 최종 데이터셋

- 서울시 자치구별 쓰레기 발생량, 주택수, 1인가구 평균 배달일수, 1인가구 비율

지역	쓰레기 발생량	주택수	1인가구 평균 배달일수	1인가구 비율
종로구	264.2	30,259	12.687895	0.162980
중구	357.5	¹⁴ 15,263	13.326083	0.168956
용산구	301.2	36,157	12.914051	0.150529
성동구	301.2	22,673	12.816761	0.140947
광진구	343.1	56,244	14.356042	0.172316

02 데이터 소개

■ 시각화

- **seaborn** 라이브러리의 **barplot** 함수 이용하여 막대 그래프 시각화
- **folium** 라이브러리의 **choropleth** 이용하여 지도 시각화

```
import seaborn as sns
plt.figure(figsize = (10,4))
plt.title("지역별 쓰레기 발생량", fontsize = 20)
sns.set_theme(style="whitegrid")
sns.barplot(data=df3, x='지역', y='쓰레기발생량', palette='RdYlGn')
plt.xticks(rotation=45)
plt.rc('axes', labelsz=15) # x,y축 label 폰트 크기
plt.rc('xtick', labelsz=13) # x축 눈금 폰트 크기
plt.show()
```

```
import folium
import requests
import json

# 서울 행정구역 json raw파일(githubcontent)
r = requests.get('https://raw.githubusercontent.com/southkorea/seoul-maps/master/kostat/2013/json/seoul_municipalities_geo_simple.json')
c = r.content
seoul_geo = json.loads(c)

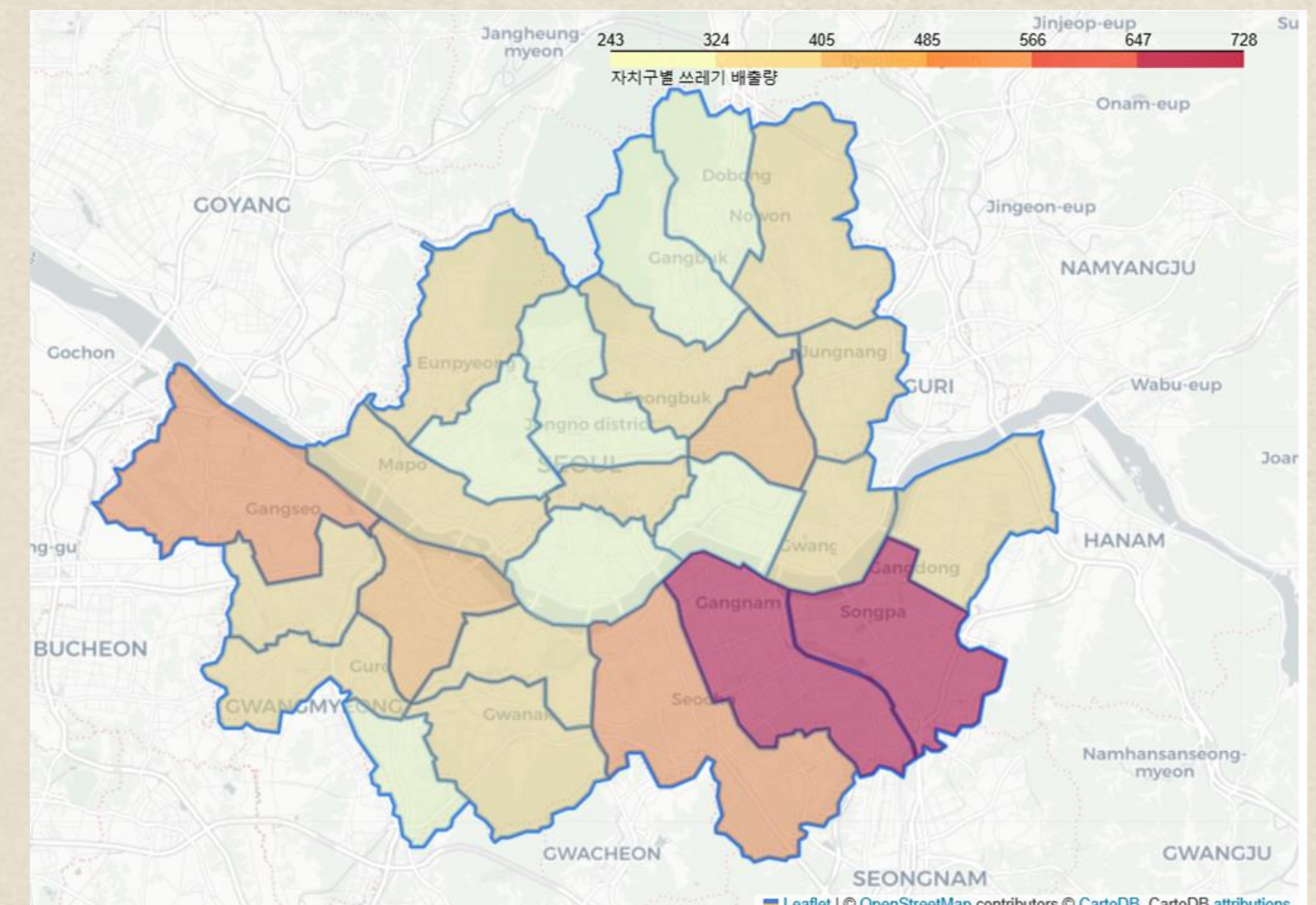
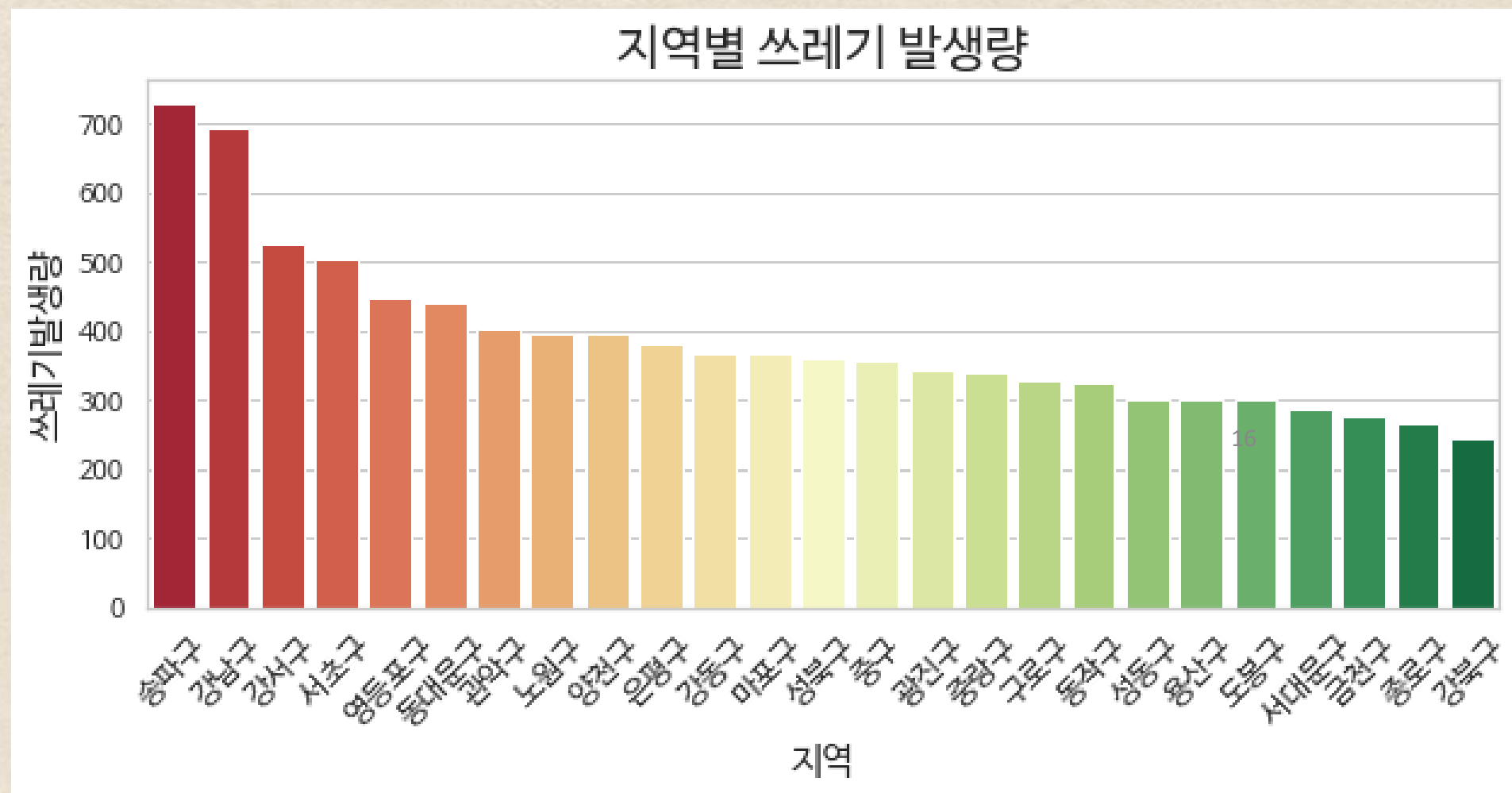
m = folium.Map(location=[37.559819, 126.963895], zoom_start=11, tiles='cartodbpositron')

folium.GeoJson(seoul_geo, name='지역').add_to(m)

m.choropleth(geo_data=seoul_geo,
              data=df3,
              columns=['지역', '쓰레기발생량'],
              fill_color='YlOrRd',
              fill_opacity=0.5,
              line_opacity=0.2,
              key_on='feature.properties.name',
              legend_name="자치구별 쓰레기 배출량"
            )
```


02 데이터 소개

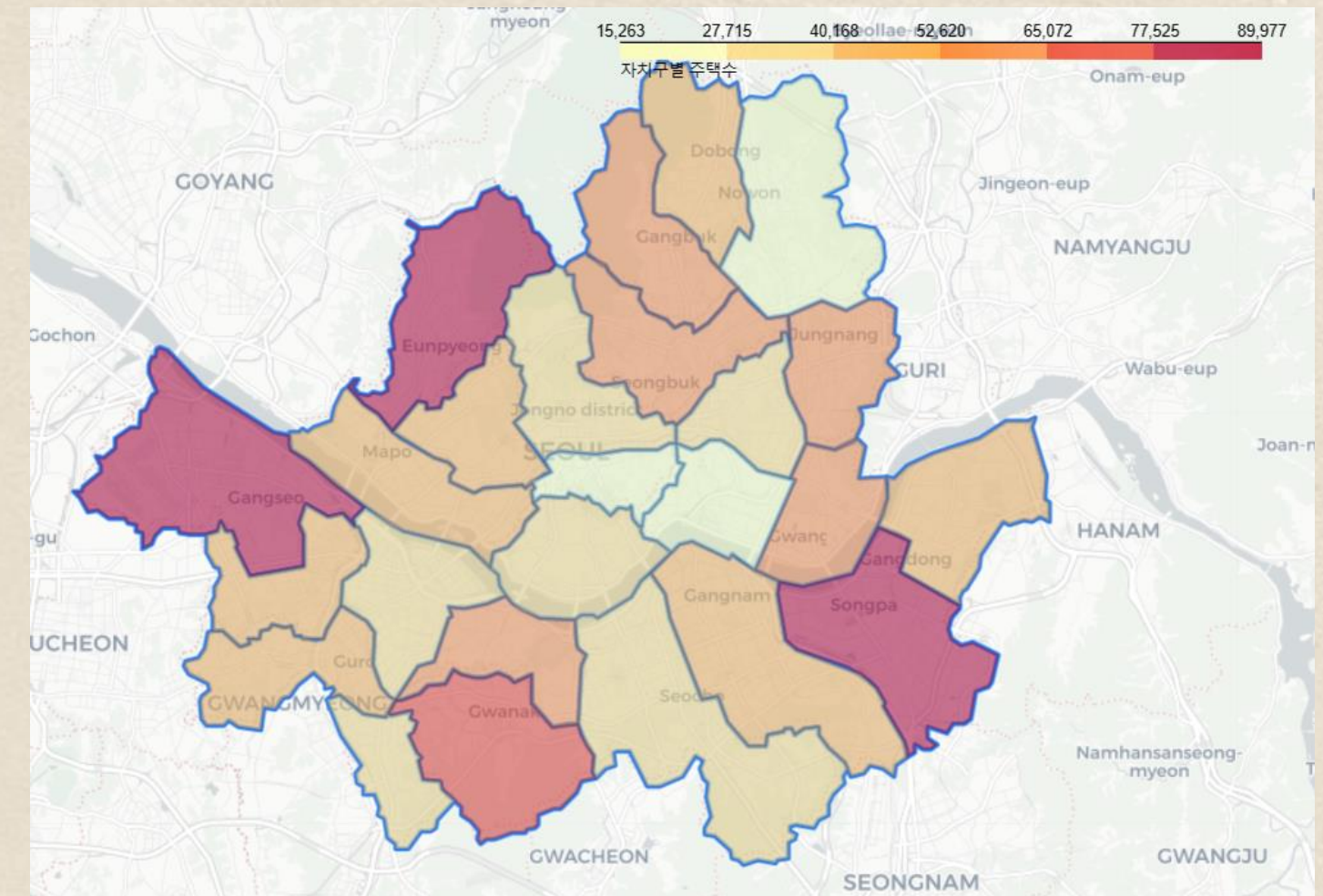
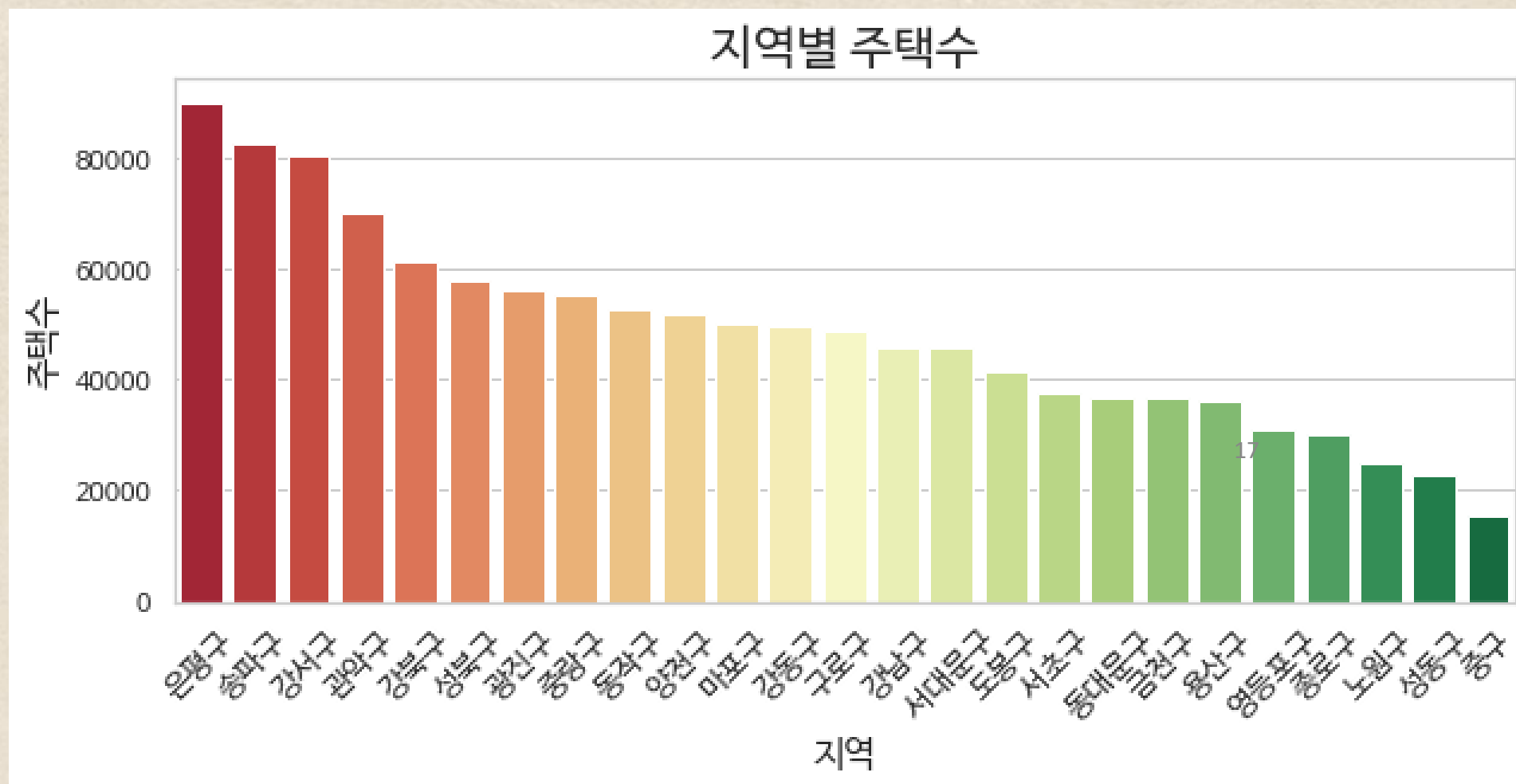
시각화



송파구, 강남구의 쓰레기 발생량이 하루 약 700톤으로 가장 높았음

02 데이터 소개

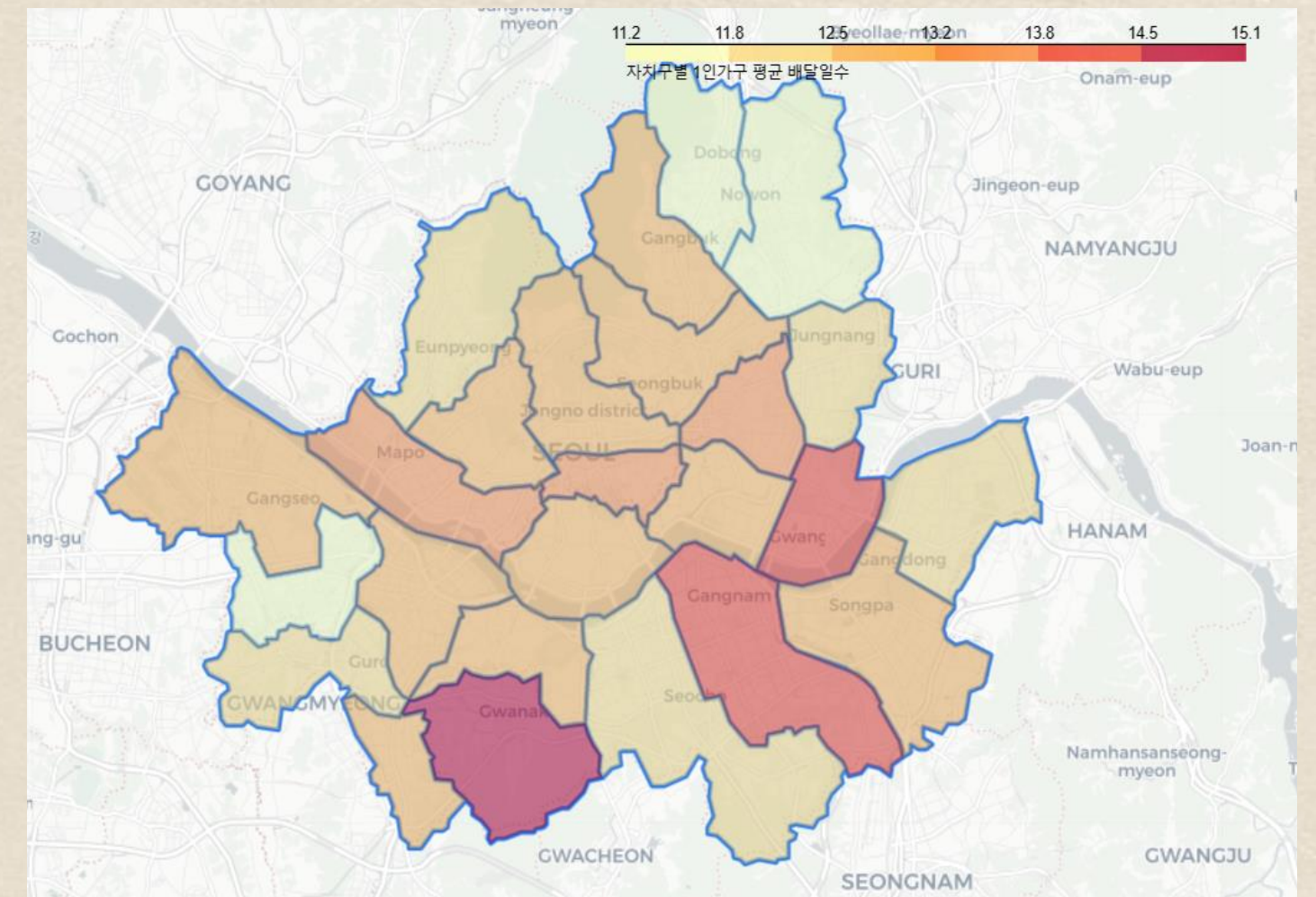
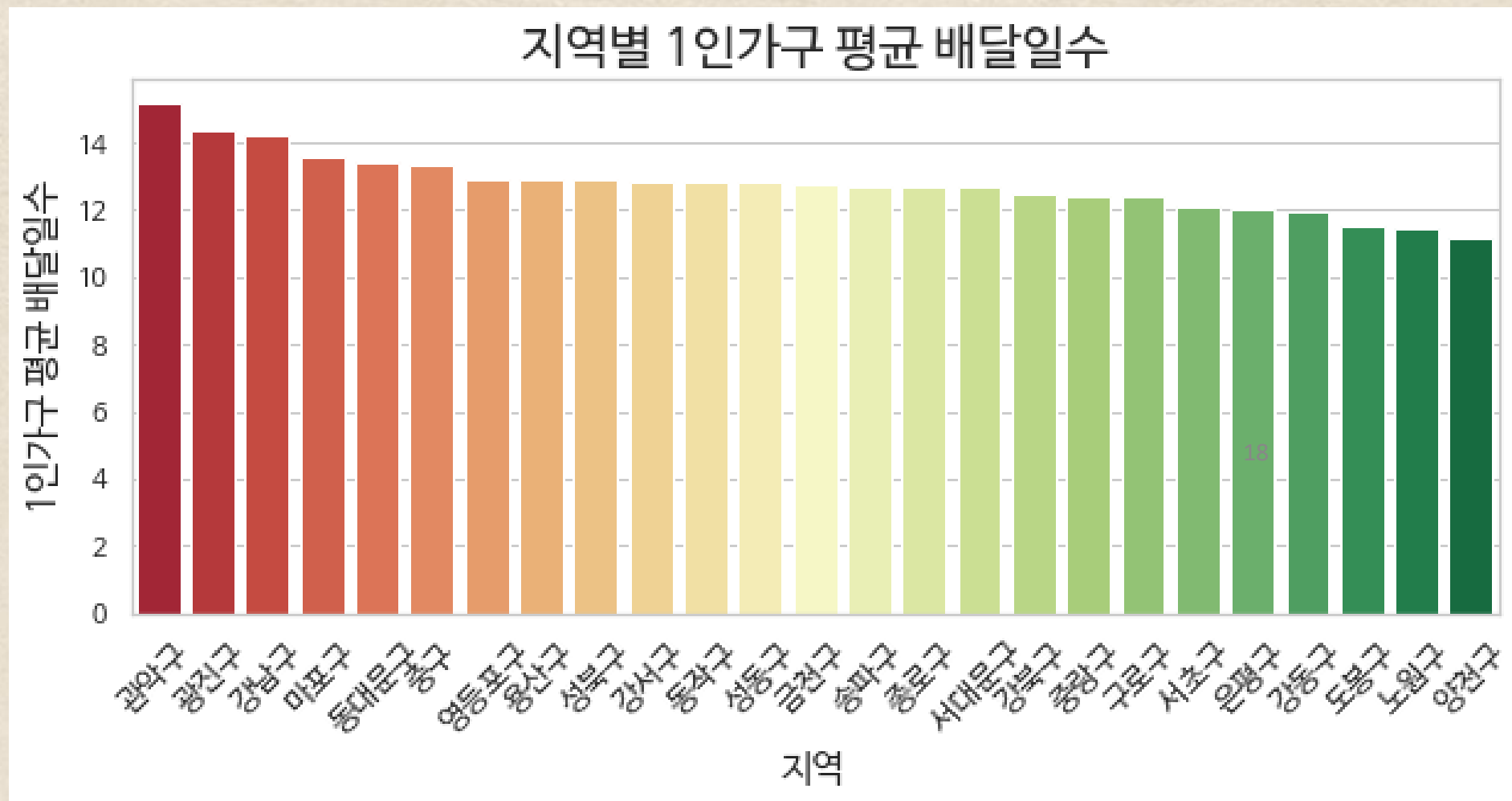
■ 시각화



은평구, 송파구, 강서구, 관악구가 주택이 가장 많았으며, 중구가 가장 적었음

02 데이터 소개

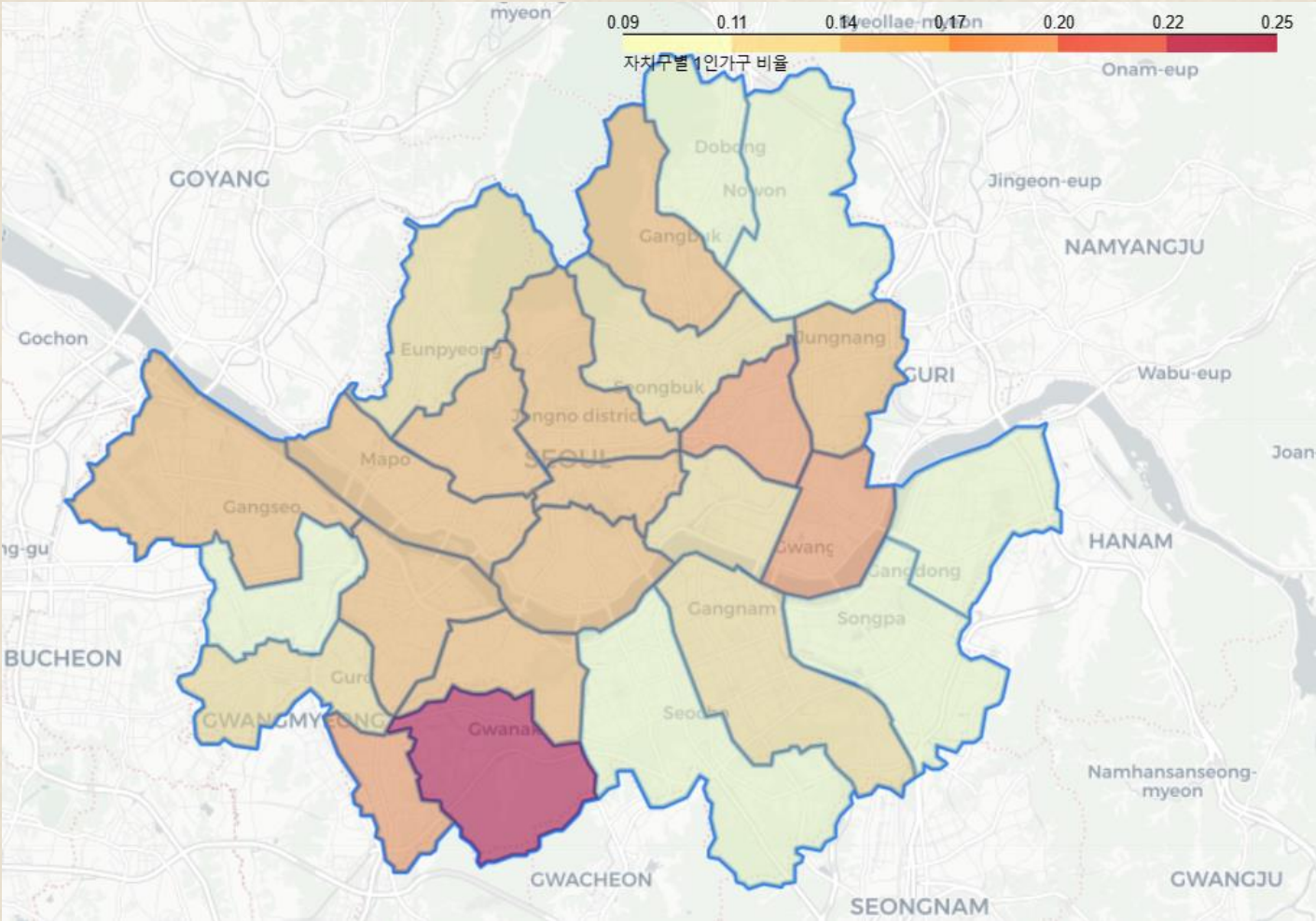
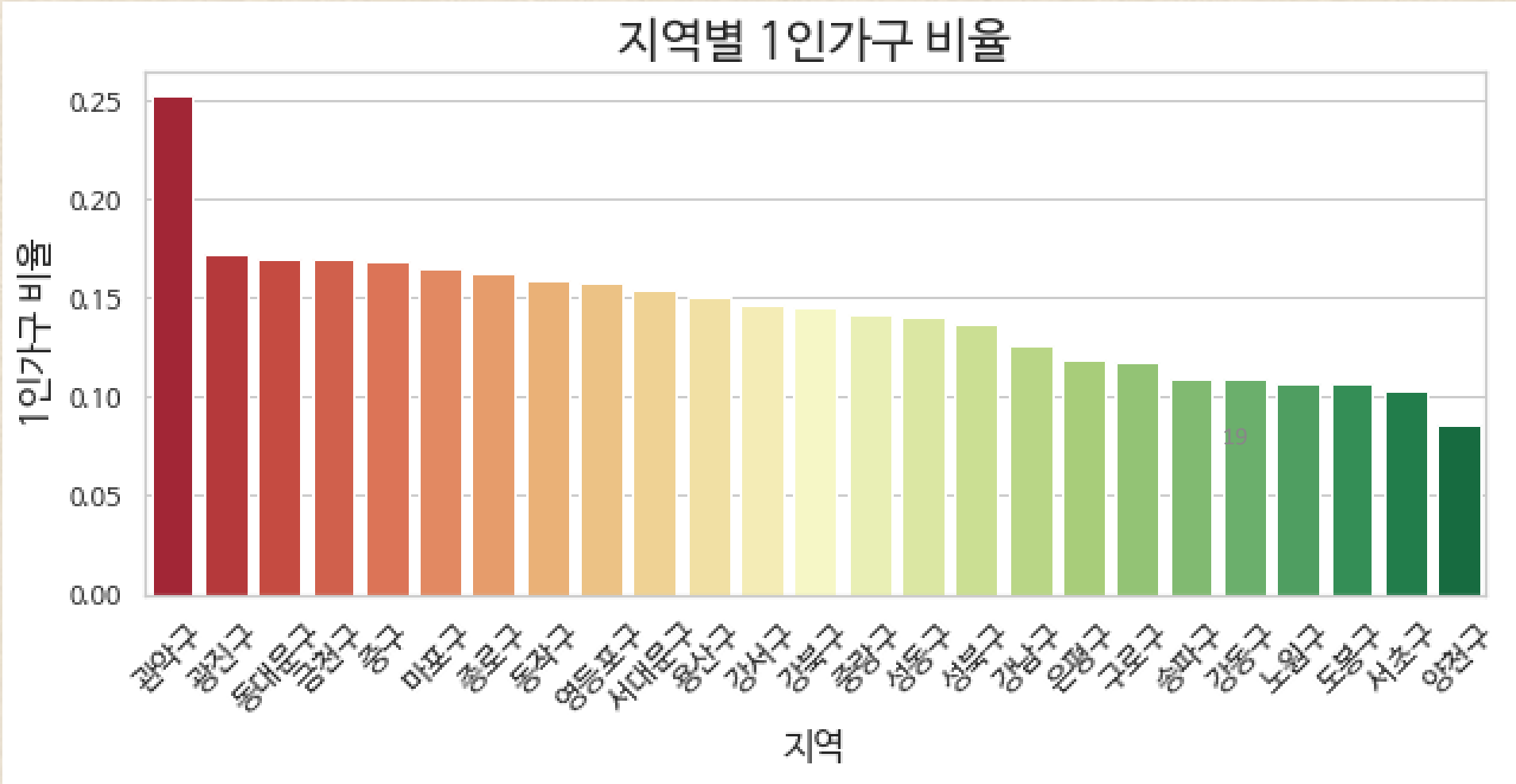
시각화



모든 자치구의 1인가구 평균 배달일수가 약 13일로 큰 편차를 보이지 않음

02 데이터 소개

시각화



관악구의 1인가구 비율이 25%로 가장 높음

03 분석 방법

03 분석 방법 | 구별 점수 산정

■ 변수 별 등급 산정

```
# 구간 10개로 나누기
df['쓰레기발생량_그룹'] = pd.cut(df['쓰레기발생량'], bins = 10, labels = list(range(1,11,1)))
```

- 각 변수들을 동일한 10개의 간격으로 나누어 그룹화

지역	쓰레기발생량	...	쓰레기발생량_그룹	주택수_그룹	1인가구 평균 배달일수_그룹	1인가구 비율_그룹
종로구	264.2	...	1	3	4	5
중구	357.5	...	3	1	6	5
용산구	301.2		2	3	5	4
관악구	401.6	...	4	8	10	10
성동구	301.2	...	2	1	5	4

03 분석 방법 | 구별 점수 산정

■ 자치구 별 점수 산정

- 변수별로 나눈 그룹을 모두 더해 자치구별 점수 산정
- 점수가 높을 수록 거점형 분리 배출 시설의 필요성이 높은 자치구



지역	쓰레기발생량_그룹	주택수_그룹	1인가구 평균 배달일수_그룹	1인가구 비율_그룹	점수
종로구	1	3	4	5	13
중구	3	1	6	5	15
용산구	2	3	5	4	14
관악구	4	8	10	10	32
성동구	2	1	5	4	12

03

분석 방법 | 관악구 주택 위치 데이터 추출

```
#1. 관악구 주택 현황 데이터 전처리
# 시군구, 법정동명, 본번, 부번 열 생성
df = pd.read_csv('./data/raw/서울특별시 관악구_법정동 주택현황_20221121.csv', encoding='cp949')
df1 = df.query("주용도 in ['단독주택', '공동주택']")
df1 = df1[['시군구', '법정동명', '본번', '부번']]
df1 = df1.drop_duplicates() # 중복행 제거

#2. 관악구 아파트 위치 데이터 전처리
# 시군구, 법정동명, 본번, 부번 열 생성
df2 = pd.read_csv('./data/raw/관악구 아파트.csv', encoding='utf-8')
df2['시군구'] = [i.split()[0] + ' ' + i.split()[1] for i in df2['지번 주소']]
df2['법정동명'] = [i.split()[2] for i in df2['지번 주소']]
df2['본번부번'] = [i.split()[3] for i in df2['지번 주소']]
df2['본번'] = [i.split('-')[0] for i in df2['본번부번']]

aa = [i.split('-') for i in df2['본번부번']]
j=[]
for i in aa:
    if len(i)==1:
        i.append('0')
    j.append(i)
df2['부번']=[i[1] for i in j]

#3. 관악구 주택 데이터에서 아파트 필터링
df3 = pd.concat([df1, df2[['시군구', '법정동명', '본번', '부번']]], axis=0 )
df4 = df3.drop_duplicates()
df4.to_csv('./주택데이터정리.csv', encoding='cp949', index=False)
```


03 분석 방법 | 관악구 주택 위치 데이터 추출

■ 관악구 주택 현황 데이터

- 공공데이터 포털에서 제공하는 「서울특별시 관악구_법정동 주택현황」
- 2022년 관악구 주택 주소(시군구, 법정동명, 본번, 부번), 주용도 데이터 이용
- 주용도가 공동주택, 단독주택인 데이터만 활용

시군구	법정동명	본번	부번	주용도
서울특별시 관악구	봉천동	942	7	단독주택
서울특별시 관악구	봉천동	100	172	단독주택
...
서울특별시 관악구	봉천동	1588	21	공동주택
서울특별시 관악구	신림동	1426	10	공동주택

출처: <https://www.data.go.kr/data/15108260/fileData.do>

03 분석 방법 | 관악구 주택 위치 데이터 추출

■ 관악구 아파트 위치 데이터 추출

- 공동주택에는 아파트도 포함 → 아파트 필터링 필요
- 도로명 주소 사이트 크롤링을 통해 284개의 관악구 아파트 주소 추출 후 제거

지번 주소	본번	부번
서울특별시 관악구 봉천동 1701 건영아파트(6차) <small>25</small>	1701	0
서울특별시 관악구 봉천동 1000 현대아파트	1000	0
서울특별시 관악구 신림동 1709 양씨아파트	1709	0
서울특별시 관악구 신림동 746-43 건영아파트	746	43
서울특별시 관악구 신림동 1700 라이프아파트	1700	0

03

분석 방법 | 관악구 주택 위치 데이터 추출

최종 데이터셋

- 23,454 행의 관악구 주택 위치 데이터

시군구	법정동명	본번	부번
서울특별시 관악구	봉천동	942	7
서울특별시 관악구	봉천동	100	172
...
서울특별시 관악구	봉천동	100	519
서울특별시 관악구	봉천동	1728	0

03 분석 방법 | Geocoding

Geocoding Tool을 활용한 주택 위도, 경도 데이터 추출

_시군구	법정동명	본번	부번
서울특별시 관악구	봉천동	942	7
서울특별시 관악구	봉천동	100	172
서울특별시 관악구	봉천동	1609	20
서울특별시 관악구	신림동	593	0
서울특별시 관악구	봉천동	1688	112
서울특별시 관악구	신림동	1450	8

주소칼럼 선택 27

주소 문자열 생성에 사용할 칼럼 개수 4 개

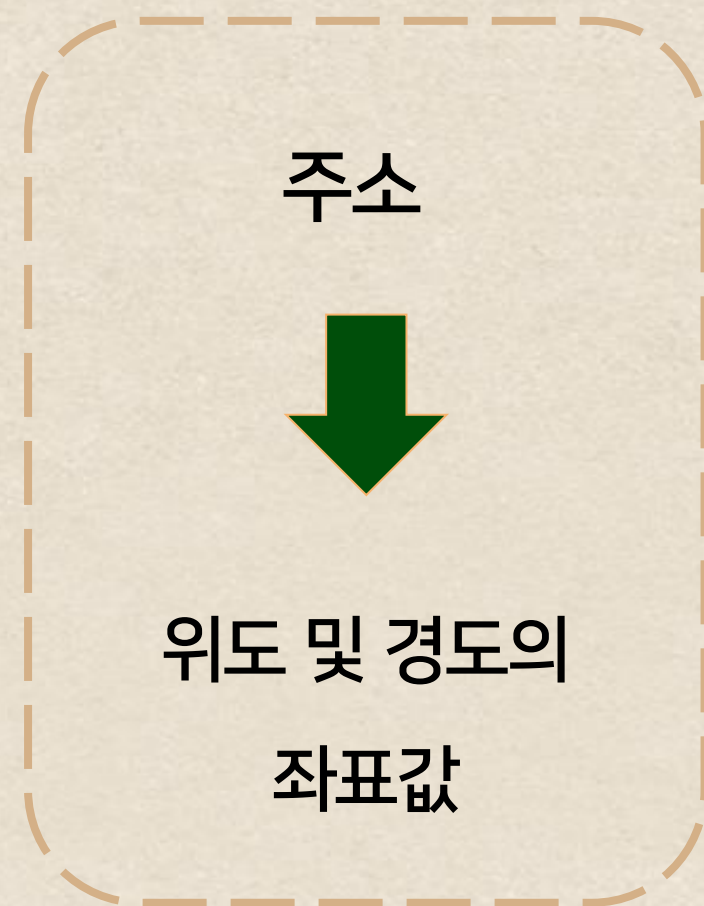
칼럼 연결문자(기본 빈칸) 칼럼 연결문자(기본 빈칸) 칼럼 연결문자(기본 빈칸)

주소 앞 부분에 추가할 문자열

위 조건을 적용한 첫 번째 주소 샘플 서울특별시 관악구 봉천동 942-7

03 분석 방법 | Geocoding

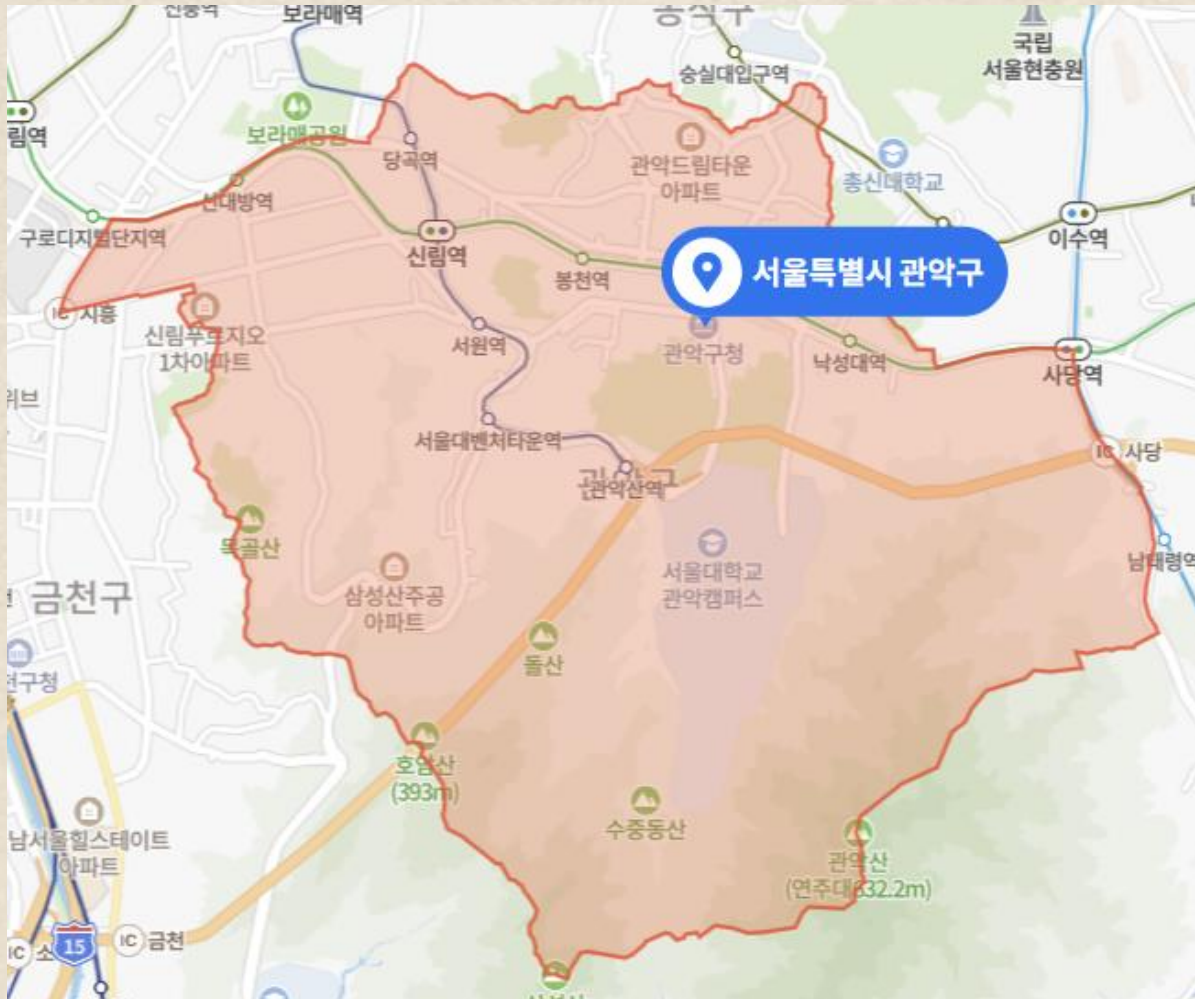
Geocoding Tool을 활용한 주택 위도, 경도 데이터 추출



입력주소	경도	위도
서울특별시 관악구 봉천동 942-7	126.940484	37.485186
서울특별시 관악구 봉천동 100-172	126.962042	37.481511
서울특별시 관악구 봉천동 1609-20	126.954527	37.477138
...
서울특별시 관악구 봉천동 100-519	126.959831	37.481907
서울특별시 관악구 봉천동 1728-0	126.961461	37.482597

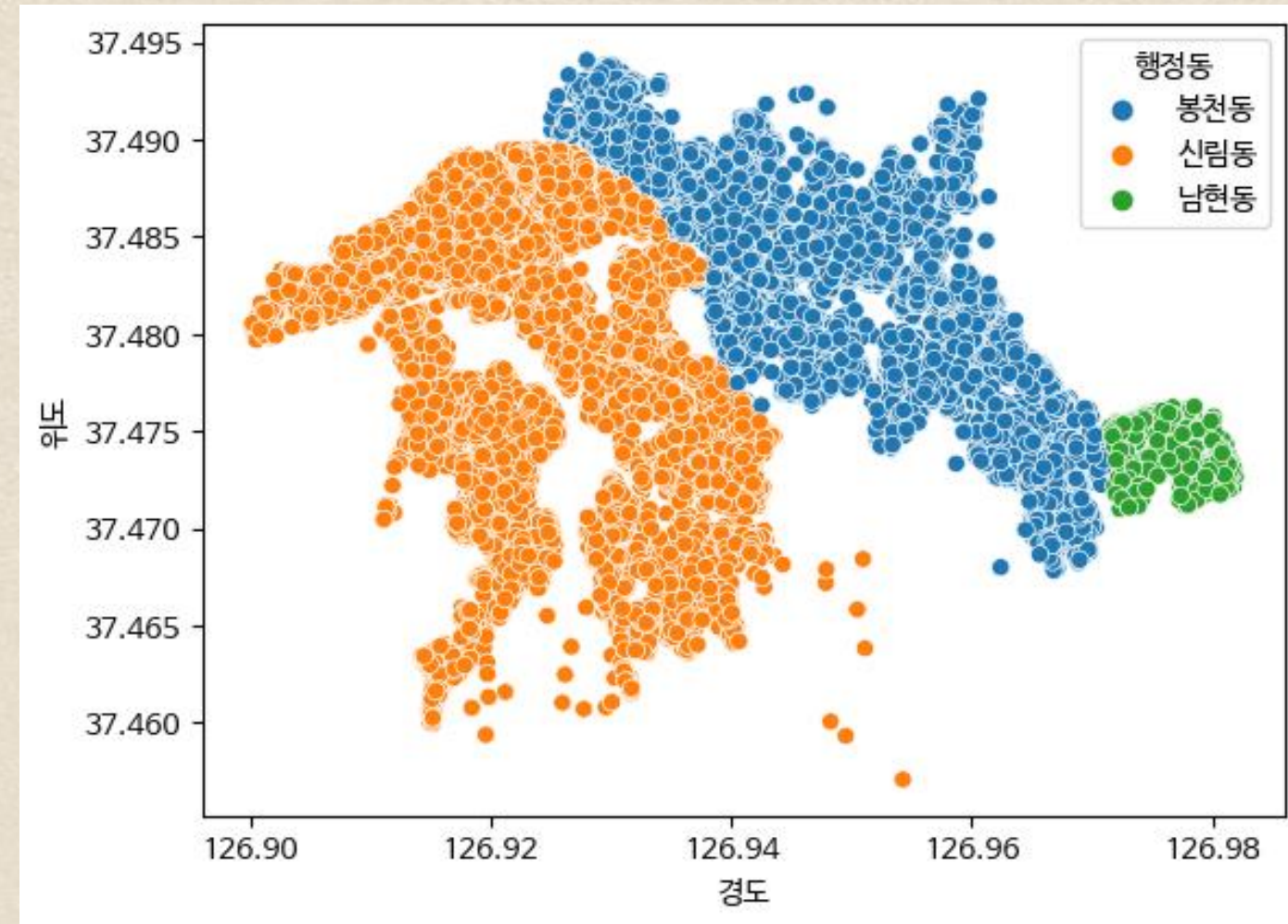
03 분석 방법 | Geocoding

관악구 주택 데이터 시각화



29

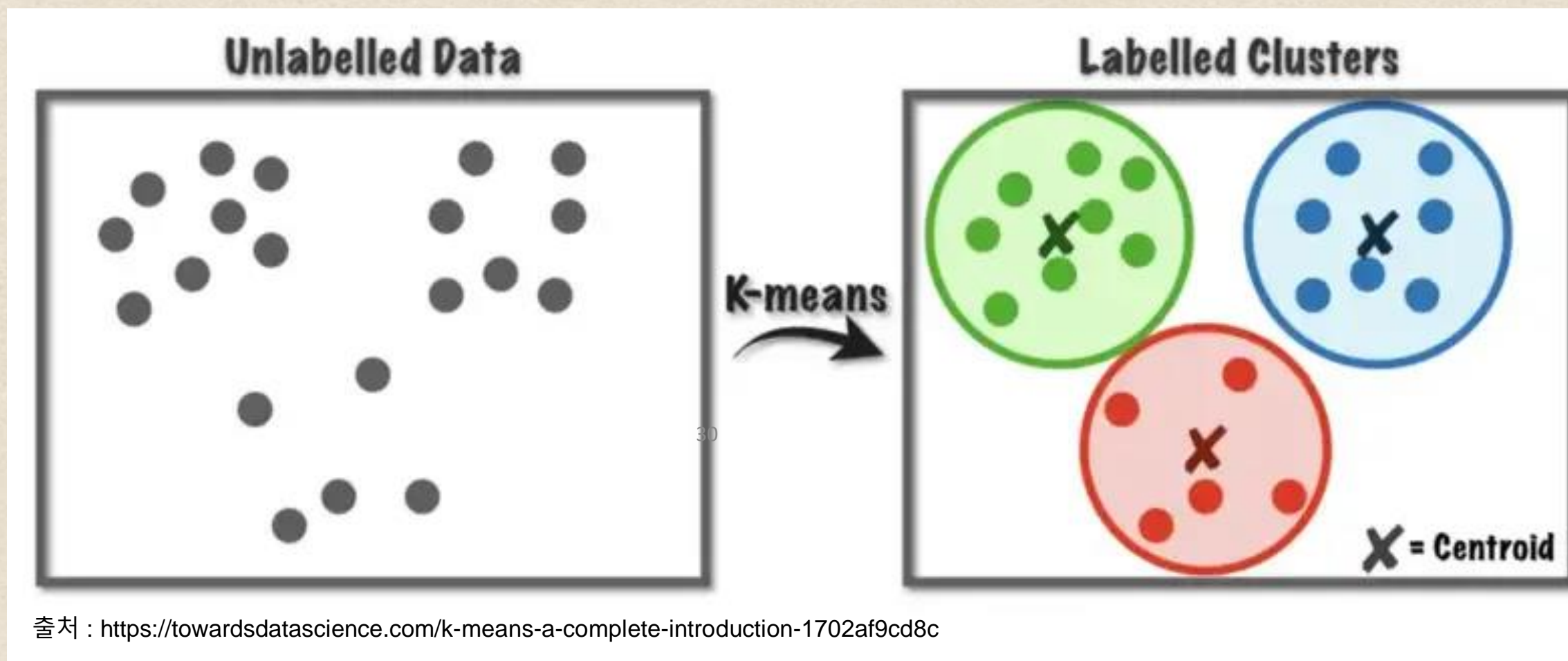
```
sns.scatterplot(data=df, x='X', y='Y', hue='행정동')  
plt.xlabel('경도')  
plt.ylabel('위도')
```



아파트, 상업시설, 산지 등을 제외한 행정동 별 주택 분포도

03 분석 방법 | K-means Clustering을 활용한 거점형 분리 배출 시설 입지 도출

K-means Clustering



- 데이터를 K개의 클러스터로 묶는 비지도 학습 알고리즘
- 군집 내 거리 최소화, 군집 간 거리 최대화 하는 방향으로 학습

03

분석 방법 | K-means Clustering을 활용한 거점형 분리 배출 시설 입지 도출

■ K-means Clustering

K-means Clustering 알고리즘을 활용해 각 그룹의 **중심**을 도출해서
해당 장소에 거점형 분리배출 시설을 설치한다면?

31

Q. 분리배출 시설 개수(K)는 어떻게 정하지?

A. 모든 주택에서 반경 100M 이내에 분리배출 시설이 위치하도록 만들자!

(해외 사례) 체코의 경우 도보 2분 거리 (90m) 에 거점형 분리수거장을 배치하여 접근성을 높임.

→ $k = 200$ 으로 설정

03

분석 방법

| K-means Clustering을 활용한 거점형 분리 배출 시설 입지 도출

```
# 클러스터링
from sklearn.cluster import KMeans
k = 200
model = KMeans(n_clusters = k, random_state = 1234)
model.fit(df[['X', 'Y']])
df['cluster'] = model.fit_predict(df[['X', 'Y']])
df_center = pd.DataFrame(model.cluster_centers_)

# 지도 시각화
map = folium.Map(location=[df_center.loc[:,1].mean(), df_center.loc[:,0].mean()], zoom_start=11)

for n in df_center.index:
    location = [df_center.loc[n, 1], df_center.loc[n, 0]]
    folium.Marker(
        location = location,
    ).add_to(map)

for n in df_center.index:
    folium.Circle(
        location=[df_center.loc[n,1], df_center.loc[n,0]],
        #icon=folium.Icon(color='red', icon='ok')
        radius = 100
    ).add_to(map)

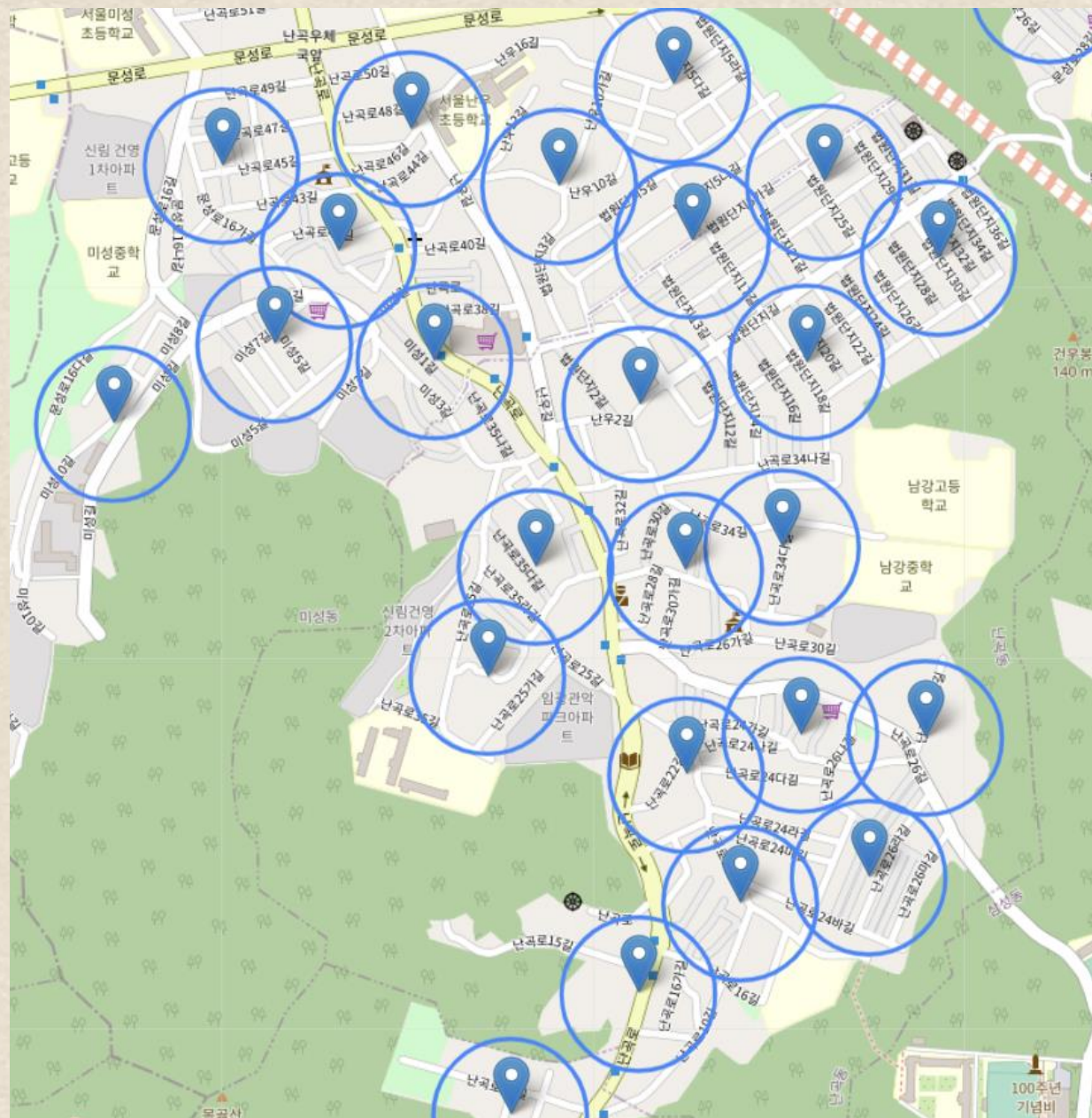
map.save('./map_cluster_final.html')
```


03

분석 방법

| K-means Clustering을 활용한 거점형 분리 배출 시설 입지 도출

K-means Clustering 결과 시각화



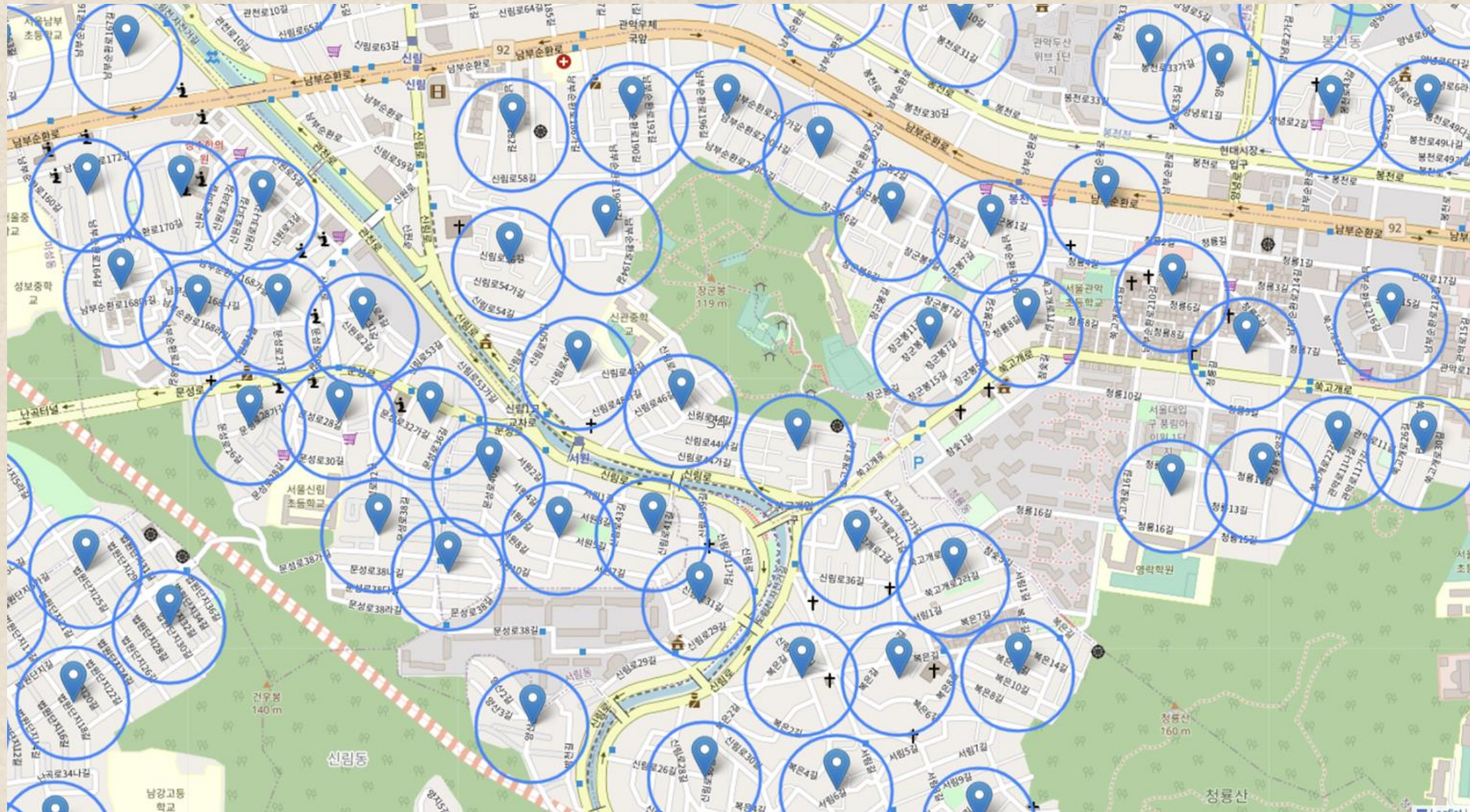
- 관악구 주택 데이터를 200 개로 군집화
- 군집화된 그룹의 중심 좌표 도출
- ³³ 각 그룹의 중심 = 거점형 분리배출 시설
- 대부분의 주택이 분리 배출 시설 100M 반경 안에 포함됨

03

분석 방법

| K-means Clustering을 활용한 거점형 분리 배출 시설 입지 도출

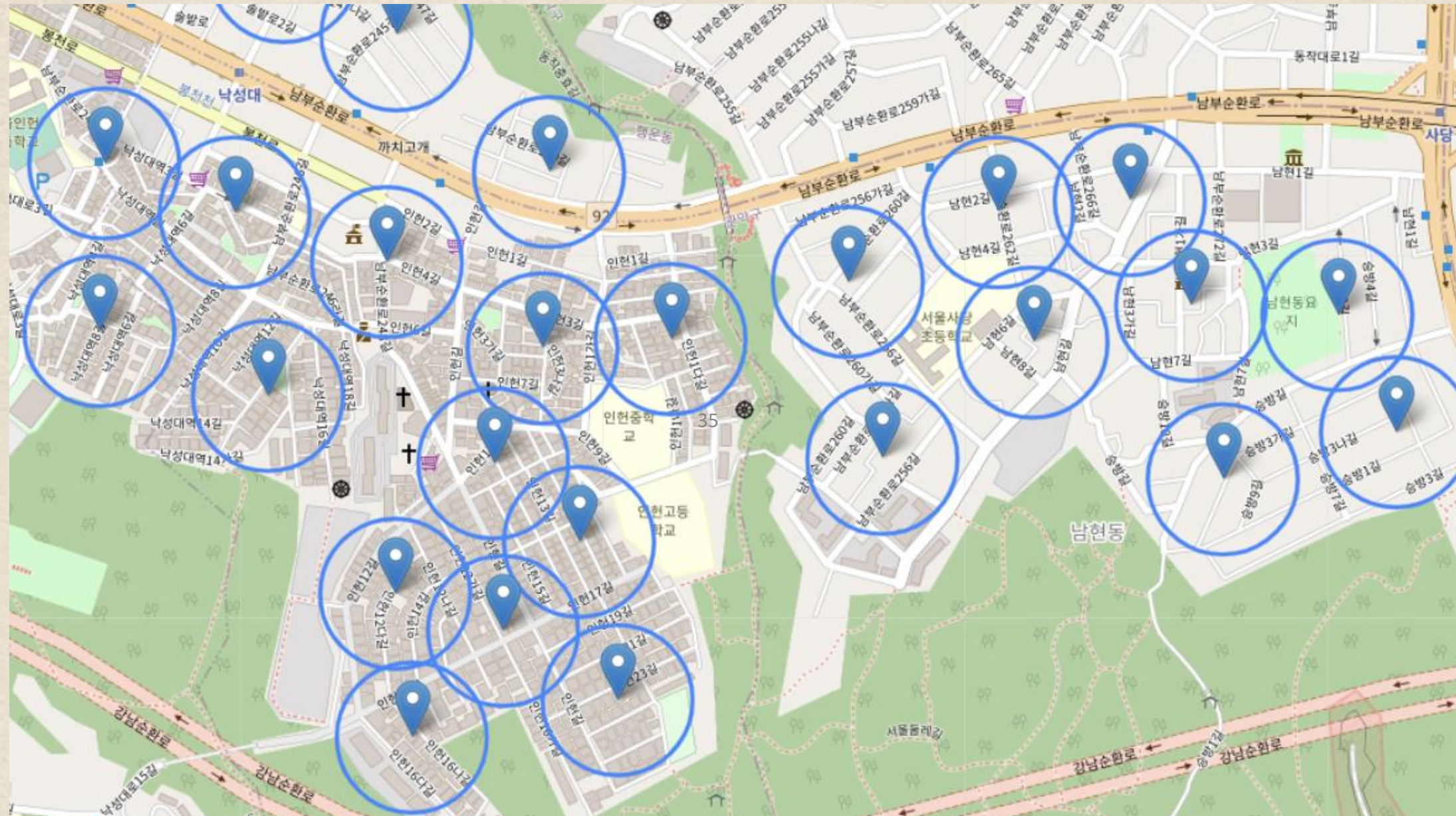
K-means Clustering 결과 시각화



분석 방법

분석 방법 | K-means Clustering을 활용한 거점형 분리 배출 시설 입지 도출

■ K-means Clustering 결과 시각화

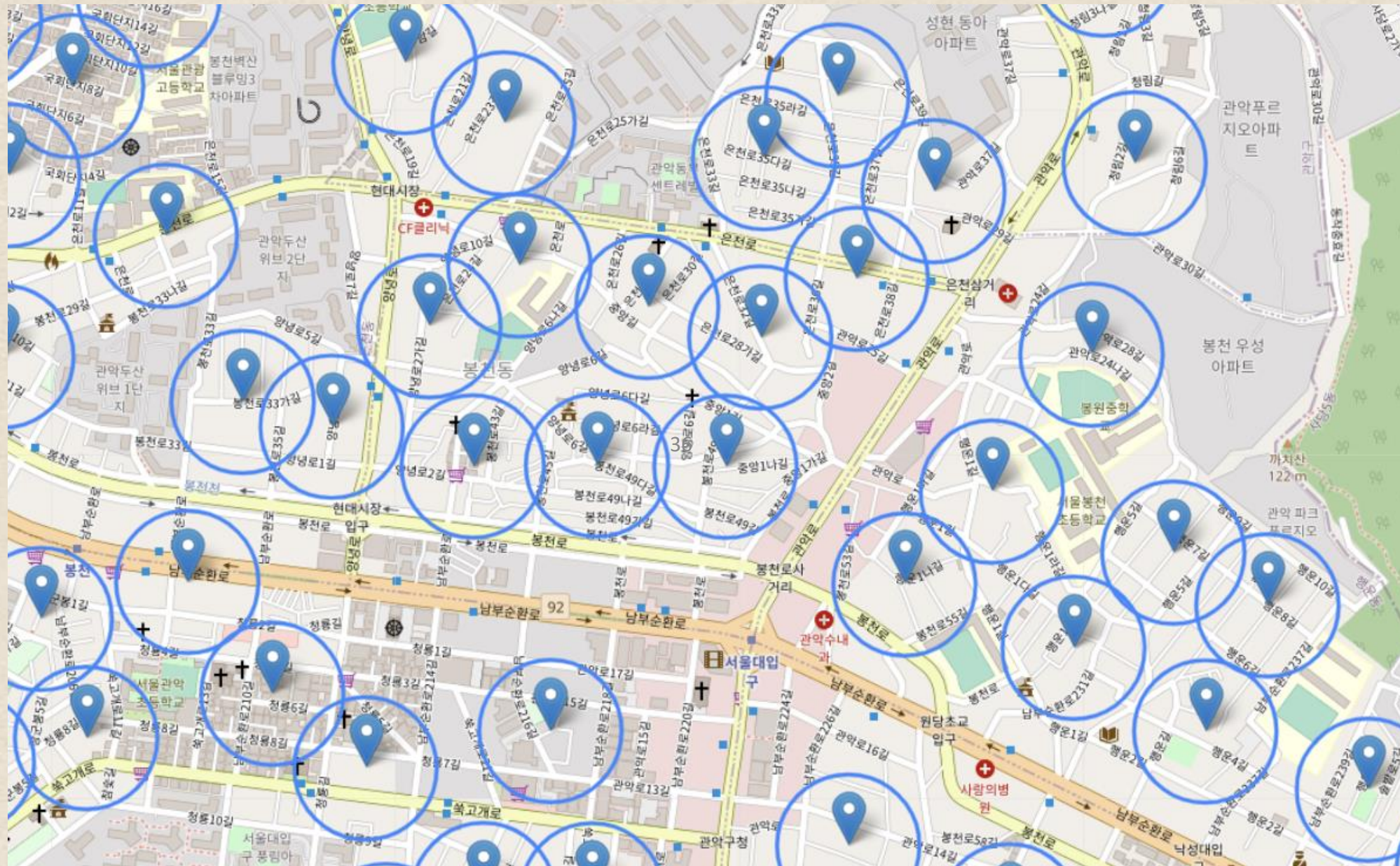


03

분석 방법

| K-means Clustering을 활용한 거점형 분리 배출 시설 입지 도출

K-means Clustering 결과 시각화

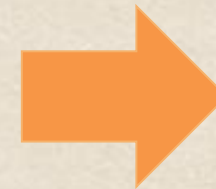


04 결론

04 결론

■ 예상 운영 시나리오

분리 수거장
설치 비용



$$174,000 \times 200 = 34,800,000$$

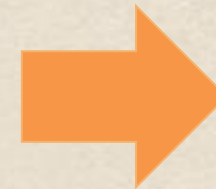
수거 기간

생활쓰레기 배출

▶ 배출시간(※토요일 제외)

· 주거지역 : 18:00부터 24:00까지

· 가로지역 : 22:00부터 익일 01:00까지



동일하게 유지 (월~금, 일)

운영 방식

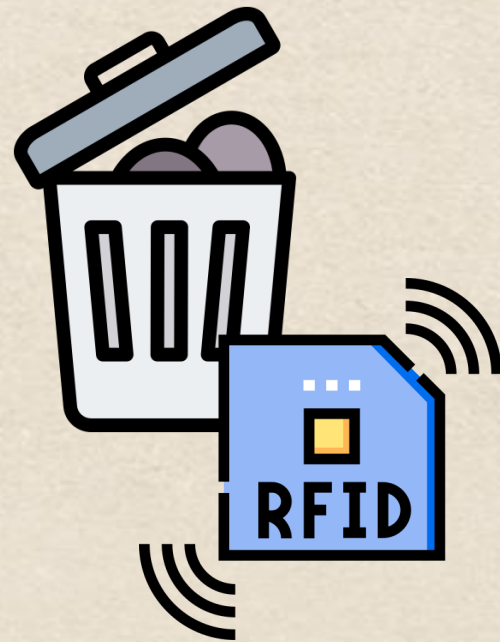
관악구 전체 주택 순회



거점형 분리 배출 시설 설치된 200곳

04 결론

■ 활용방안



RFID 기반 음식물/일반 쓰레기
종량제 기기 도입으로
쓰레기 발생 감량 효과



39

효율적인 분리수거
구역 분류로 인한
인력의 효율성 증대



분리수거 차량
최적 경로 고려 가능

04 결론

■ 기대 효과

효율적인 분리수거장 위치로 인한 무단 투기 쓰레기양 감소

길거리에 쓰레기가 없어지므로 도시 미관 개선



40



출처: '수원시 단독주택 '쓰레기감량 클린마을' 큰 성과' 기사 사진
(<https://cm.asiae.co.kr/article/2015110515165987937>)



41
감사합니다 :)