

( 광 양 시 )

# 재활용품 자동수거기 설치를 위한 최적위치 선정

5STAR(팀오성)

# CONTENTS

---

- 01** 개요
- 02** EDA / 전처리
- 03** 위치 선정을 위한 모델링
- 04** 군집화
- 05** 결론
- 06** 요약 및 한계

01

개요

- 01. 목표
- 02. 현황 파악
- 03. 분석 방향

# 01 개요

01 목표

02 현황 파악

03 분석 방향

## 공동주택 내 설치



분리수거 자동수거기,  
**공동 주택 내 설치 !!**



# 01 개요

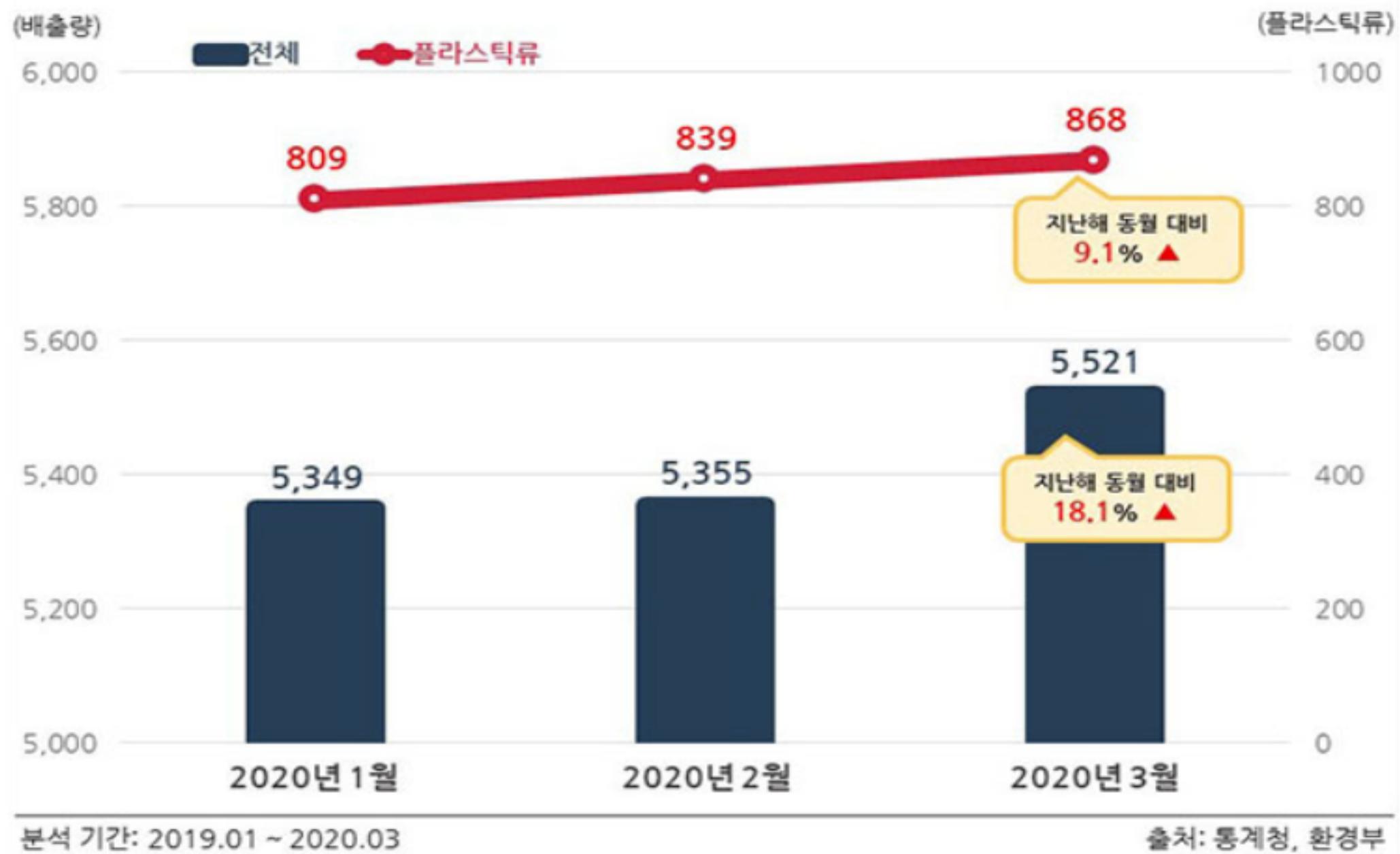
01 목표

## 02 현황 파악

03 분석 방향

# 코로나 시대, 쓰레기 대란(전국)

코로나19 이후 생활폐기물 재활용 쓰레기 품목별 통계 (단위: 톤/일)



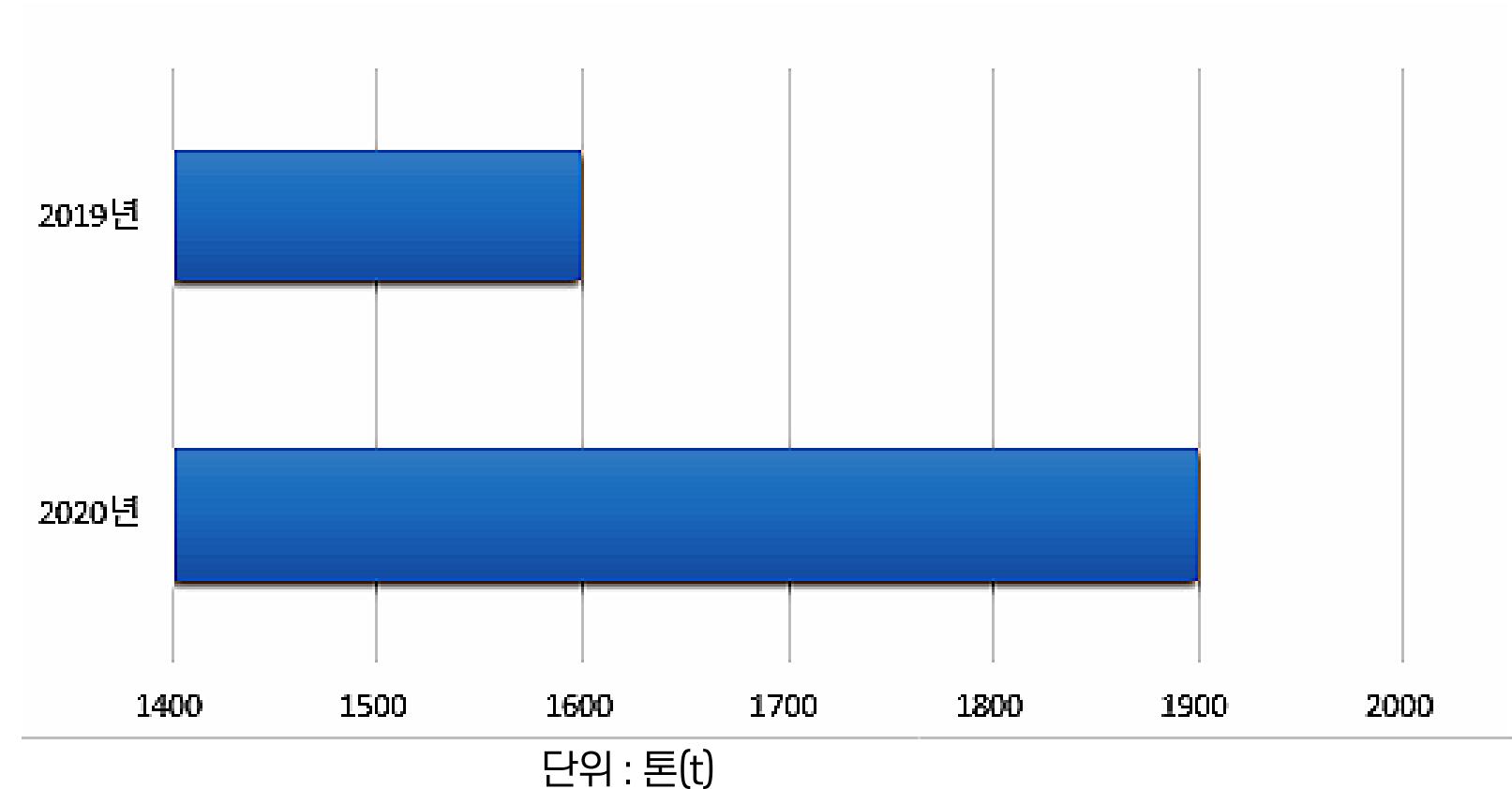
# 01 개요

01 목표

## 02 현황 파악

03 분석 방향

# 코로나 시대, 쓰레기 대란(광양)



코로나 이후  
광양시 택배 물량의 변화

“재활용 쓰레기 수거량  
전년 상반기 대비 17% 증가”

# 01 개요

01 목표

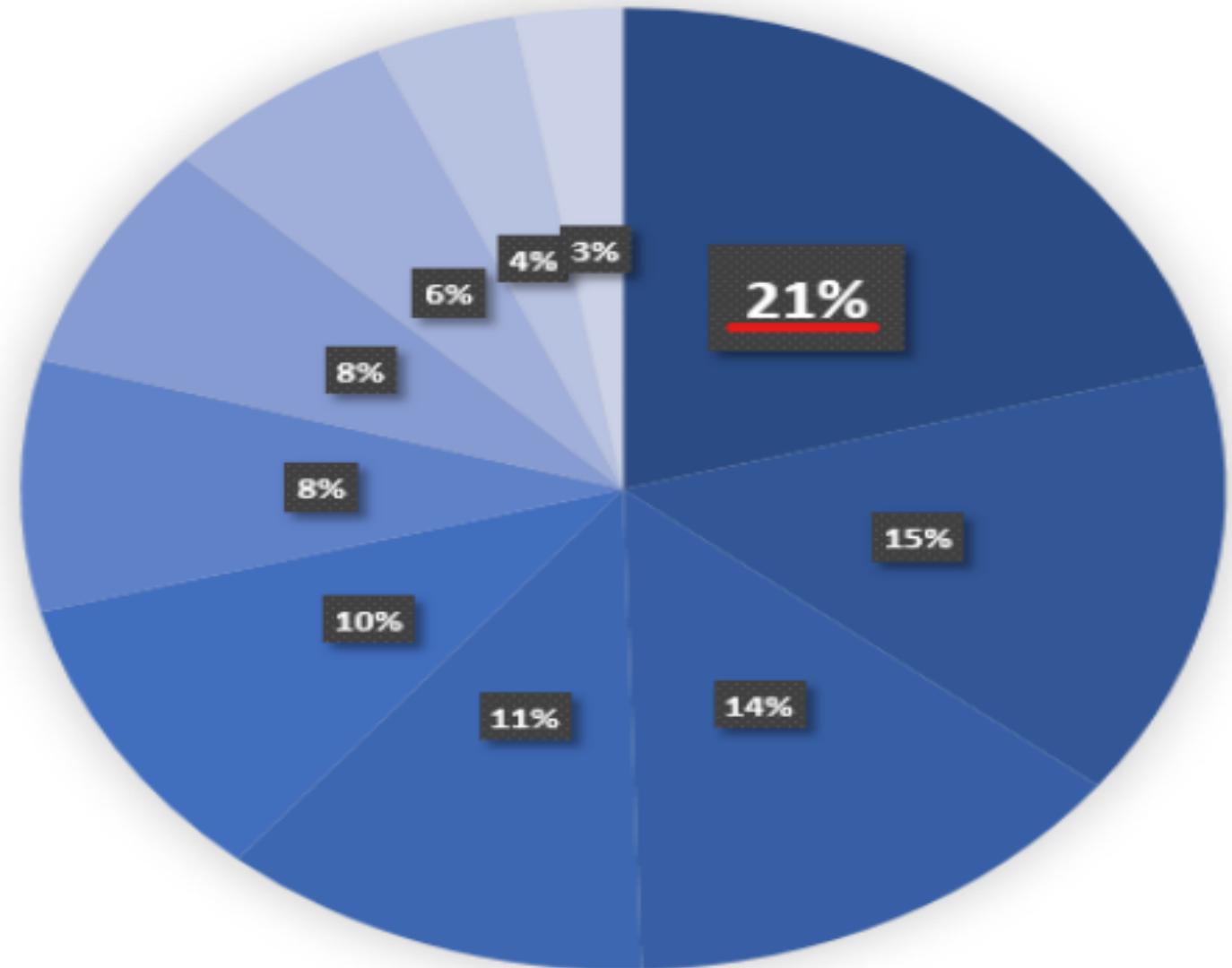
## 02 현황 파악

03 분석 방향

# 광양시 정책 및 시민의식

## 광양시민들의 요구 → **재활용 정책**

- 폐기물 감량 및 재활용 정책
  - 녹색교통정책
  - 신재생 에너지 보급 확대
  - 녹지확충 및 도시개발 제한
  - 에너지 절약
  - 기후변화 대응을 위한 전담조직 신설
  - 오염물질 총량관리
  - 온실가스 에너지 목표제 시행
  - 교통 수요관리
  - 태양광 발전사업



기후변화를 위해서 광양시가 중점적으로 추진해야 할 정책

광양시 시민 설문 조사(2019)

# 01 개요

01 목표

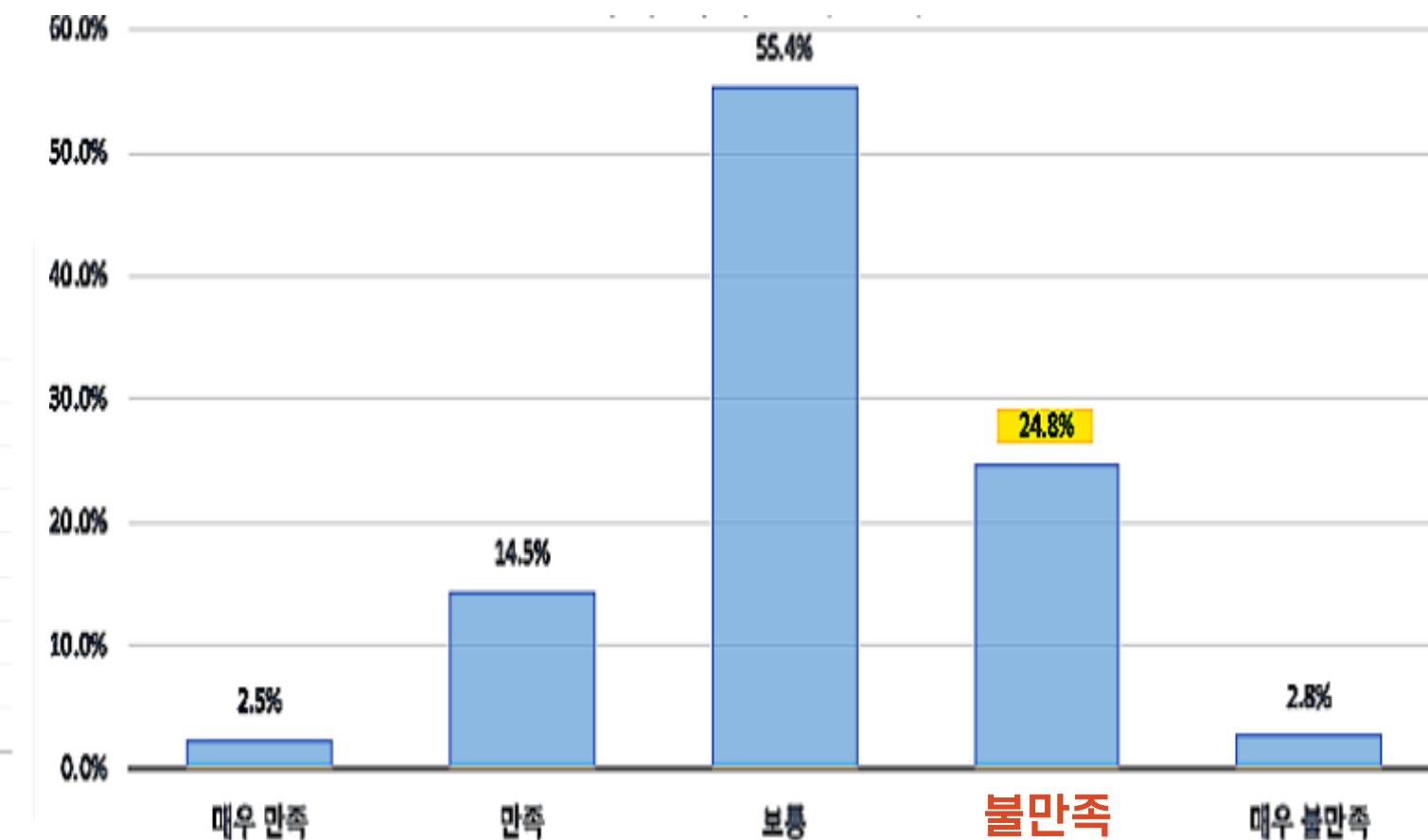
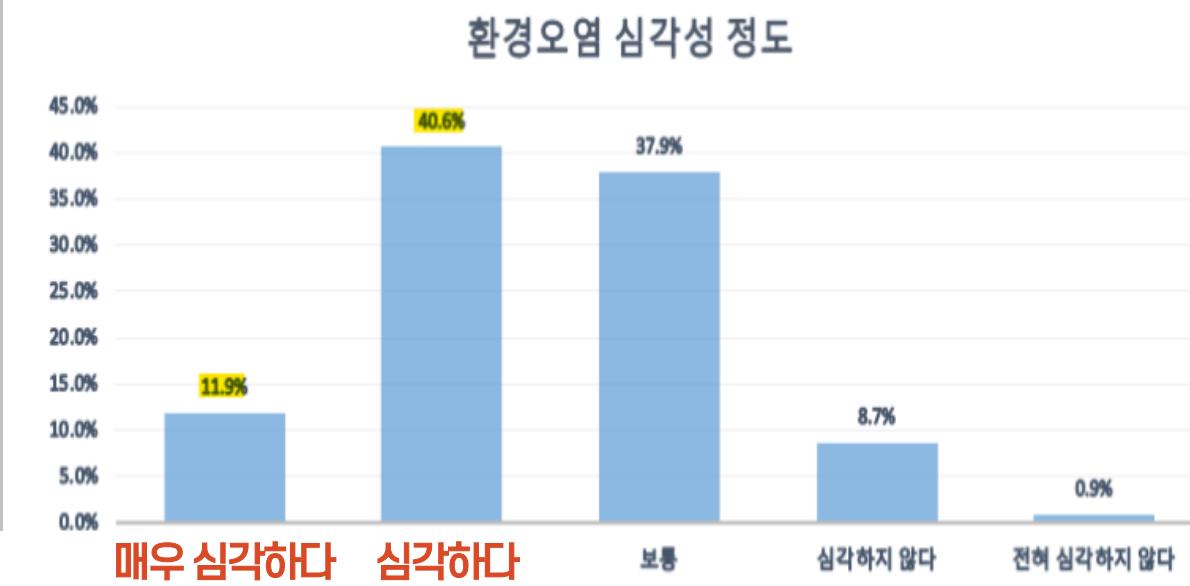
02 현황 파악

03 분석 방향

## 광양시 정책 및 시민의식

쓰레기 처리 실태에 대한 만족도를 조사한 결과

불만족 > 만족



광양시 시민 설문 조사(2019)

→ 자동수거기 설치의 이익은 광양시민에게 돌아갈 것으로 기대

# 01 개요

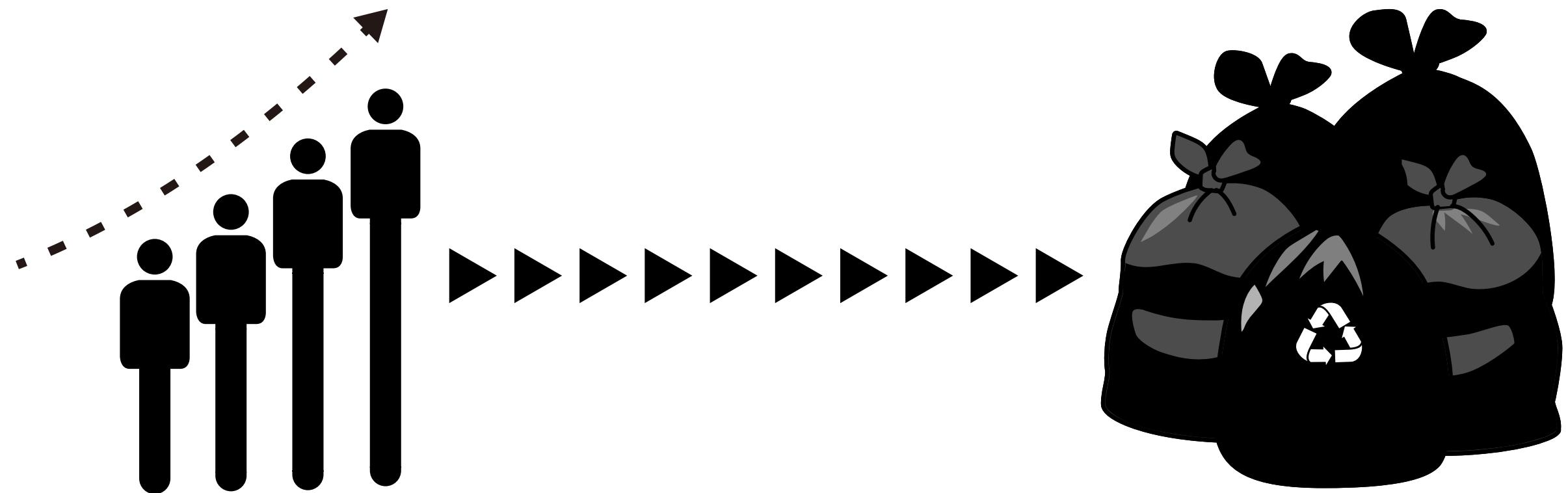
01 목표

02 현황 파악

03 분석 방향

## 광양시 정책 및 시민의식

쓰레기를 만드는 주체는 결국 사람이고,  
인구 수가 변함에 따라 쓰레기 발생량이  
**증가**하지 않을까?



# 01 개요

01 목표

02 현황 파악

03 분석 방향

## 쓰레기 발생량 ∝ 인구 수

- 생활 쓰레기 발생량은

인구 수에 비례하게 증가하는 추세를 보임.



1인가구 생활쓰레기(종류별) 발생량 조사(총량) (단위:g)

# 01 개요

01 목표

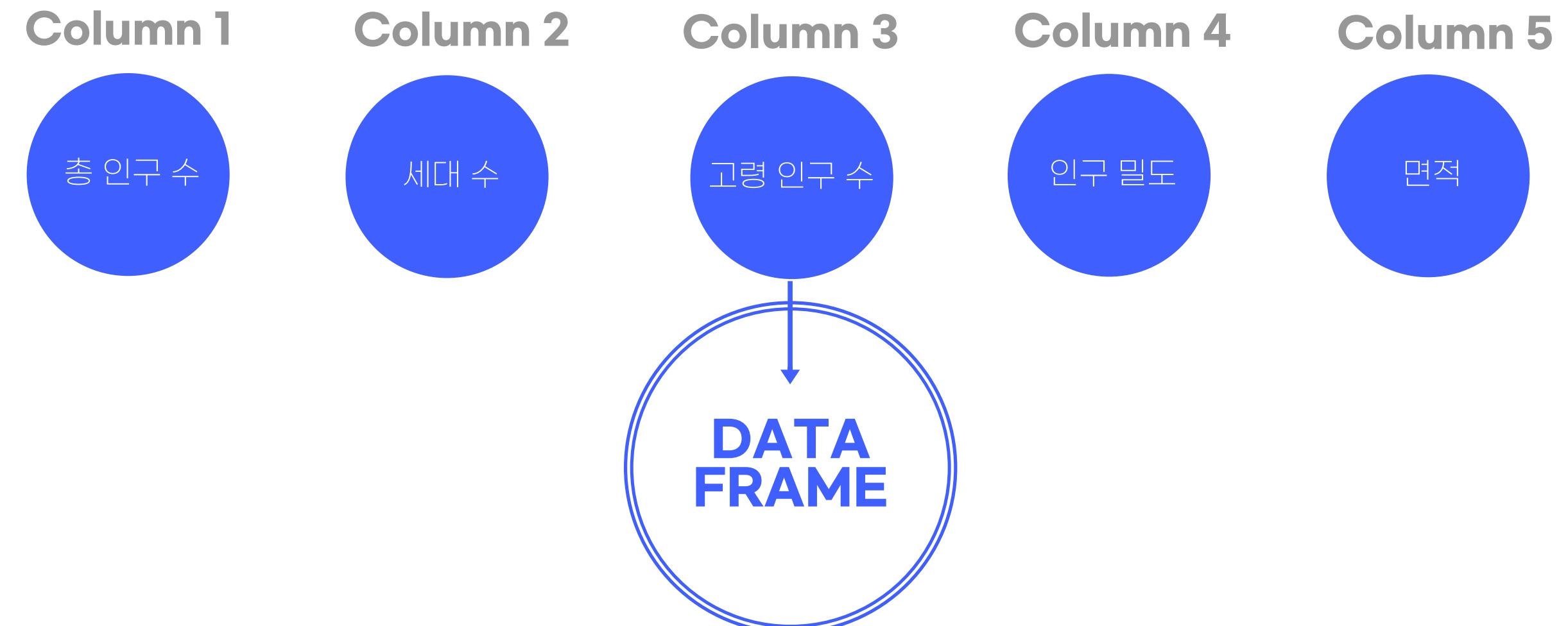
02 현황 파악

## 03 분석 방향

## 분석 방향

### 1. 인구수를 내림차순으로 1차 분류

- 신뢰성 있는 데이터 구축을 위해 세분화한 컬럼 추가



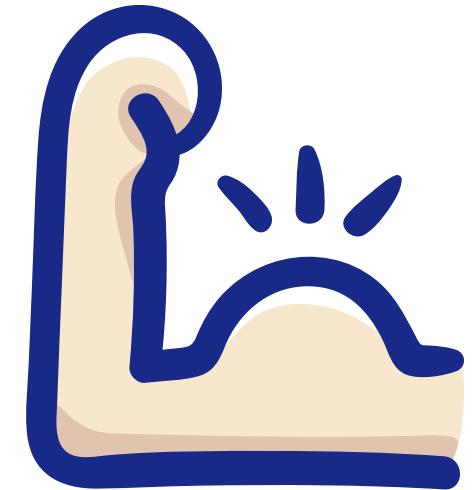
# 01 개요

01 목표

02 현황 파악

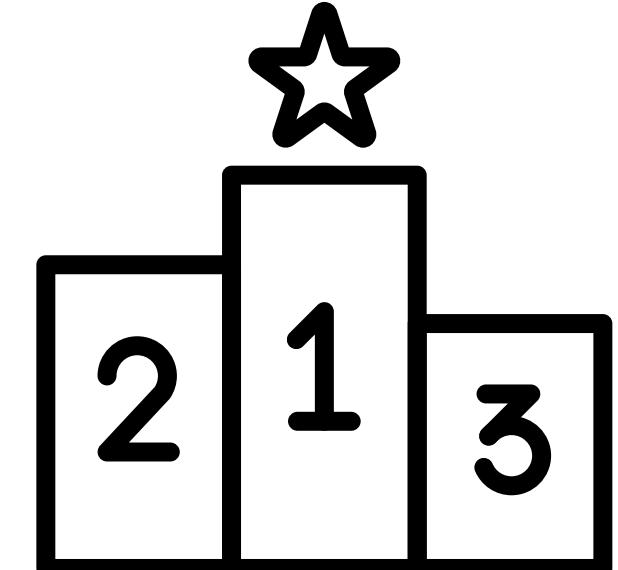
## 03 분석 방향

## 분석 방향



### 2. 최적의 수거기 설치 위치를 위한 **가중치 부여**

→ 제공받은 공개 / 비공개 데이터 여부에 따른 가중치 요소 구분  
공개(세대수) / 비공개(세대수, 분리수거장, 장려금)



### 3. 최종 점수가 높은 **50개** 공동주택 선정

# 01 개요

01 목표

02 현황 파악

## 03 분석 방향

# 분석 순서

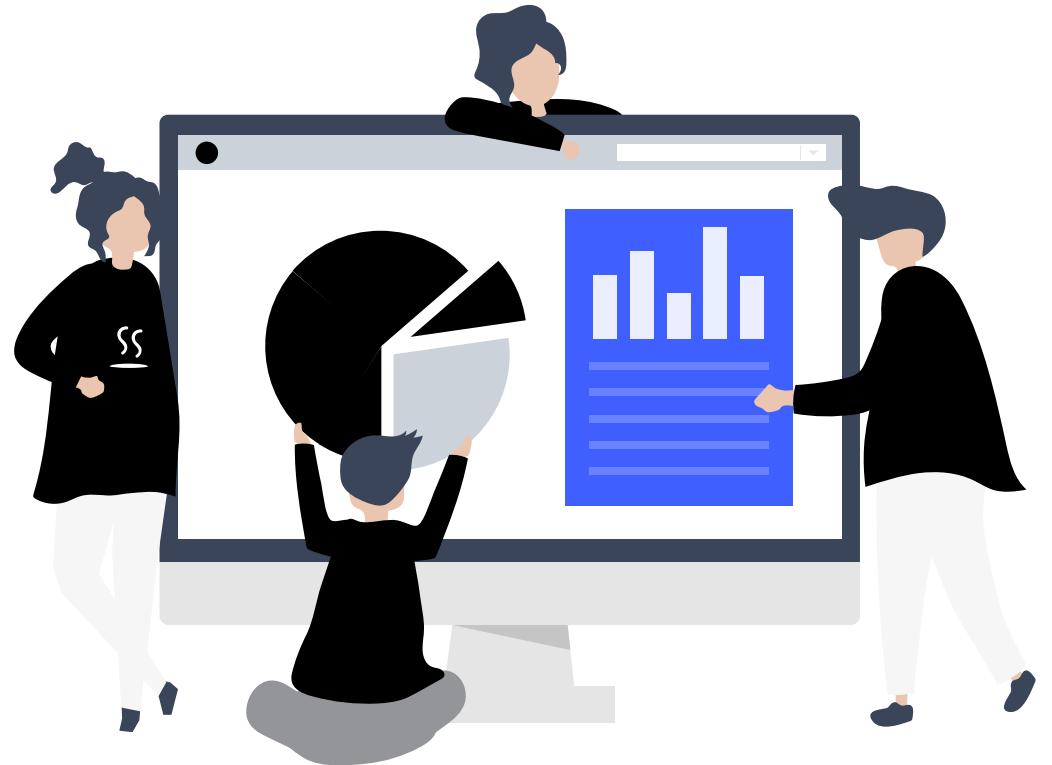


인구 수에 관한 데이터 수집 & 분석

상호 연관성 파악

미래 쓰레기 발생량 예측

위치 선정



02

EDA / 전처리

01. 데이터 불러오기
02. 데이터 전처리

# 02 EDA / 전처리

## 01 데이터 불러오기

02 데이터 전처리

## EDA



[기존 제공 데이터와 통계청의 데이터가 상이한 부분 발견]

emd_nm	total_pop_2017	total_pop_2018
광양시	153752	154585
광양시	2018	158162
광양시	2019	158437

← COMPAS 제공 데이터

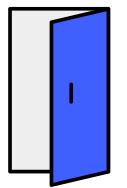
← 국가통계포털 데이터

# 02 EDA/전처리

## 01 데이터 불러오기

## 02 데이터 전처리

# 인구 / 세대수 / 고령인구



## 참고링크 및 진입 경로

The screenshot shows the KOSIS homepage with a teal header containing the KOSIS logo and the text "국내통계". Below the header, there are several navigation links: "주제별 통계", "기관별 통계", "e-지방지표(통계표)", "e-지방지표(시각화)", and "과거 · 중지통계".

This screenshot shows a search interface for regional statistics. The main title is "지역통계" (Regional Statistics). Under this, there is a list of regions: 서울특별시, 부산광역시, 대구광역시, 인천광역시, 광주광역시, 대전광역시, 울산광역시, 세종특별자치시, 경기도, 강원도, 충청북도, 충청남도, 전라북도, and 전라남도. The "전라남도" section is expanded, showing "전라남도기본통계" and "전라남도광양시기본통계". Other collapsed sections include "토지 및 기후" and "인구".



# 02 EDA/전처리

## 01 데이터 불러오기

## 02 데이터 전처리

# 인구/세대수/고령인구

▲ 전라남도

- 전라남도기본통계 [i](#)

- 전라남도광양시기본통계 [i](#)

- ▶ • 토지 및 기후

▶ 행정구역 수록기간 년 2010~2019 [o](#) [i](#)

▶ 토지지목별현황 수록기간 년 2010~2019 [o](#) [i](#)

▶ 일기일수 수록기간 년 2011~2019 [o](#) [i](#)

▶ 기상개황 수록기간 년 2011~2018 [o](#) [i](#)

▶ 강수량 수록기간 년 2010~2019 [o](#) [i](#)

▶ 해안선 및 도서 수록기간 년 2010~2019 [o](#) [i](#)

▶ 기상요소별 장기개황 수록기간 년 2011~2017 [o](#) [i](#)

- ▶ • 인구

▶ 인구추이 수록기간 년 2010~2019 [o](#) [i](#)

▶ 거소신고인수 수록기간 년 2015~2019 [o](#) [i](#)

▶ 읍면동별 세대 및 인구 수록기간 년 2010~2019 [o](#) [i](#)

▶ 연령별(5세계급) 및 성별인구 수록기간 년 2010~2018 [o](#) [i](#)

▲ 전라남도

- ▶ • 전라남도기본통계 [i](#)

- 2002년 이후

- 토지 및 기후

- 인구

▶ 인구추이 수록기간 년 2002~2019 [o](#) [i](#)

▶ 시·군별 세대 및 인구(주민등록) 수록기간 년 2002~2019 [o](#) [i](#)

▶ ▶ 읍·면·동별 세대 및 인구 수록기간 년 2002~2018 [o](#) [i](#)

- ▶ • 환경

▶ 환경오염물질 배출사업장 수록기간 년 2002~2019 [o](#) [i](#)

▶ 환경오염배출사업장 단속 및 행정조치 수록기간 년 2002~2019 [o](#) [i](#)

▶ 배출부과금 부과 및 징수현황 수록기간 년 2011~2019 [o](#) [i](#)

▶ 대기오염 수록기간 년 2005~2018 [o](#) [i](#)

▶ 쓰레기 수거 수록기간 년 2002~2017 [o](#) [i](#)

▶ 생활폐기물 매립지 수록기간 년 2002~2018 [o](#) [i](#)

▶ 폐기물재활용률 수록기간 년 2011~2018 [o](#) [i](#)

▶ 하수종말처리장 수록기간 년 2007~2017 [o](#) [i](#)

# 02

## EDA / 전처리

### 01 데이터 불러오기

02 데이터 전처리

# 인구 / 세대수 / 고령인구

국가통계포털의 데이터를 인용하여

광양시 / 행정구역별 데이터 수집

1. 세대수(세대)

2. 등록인구(명)

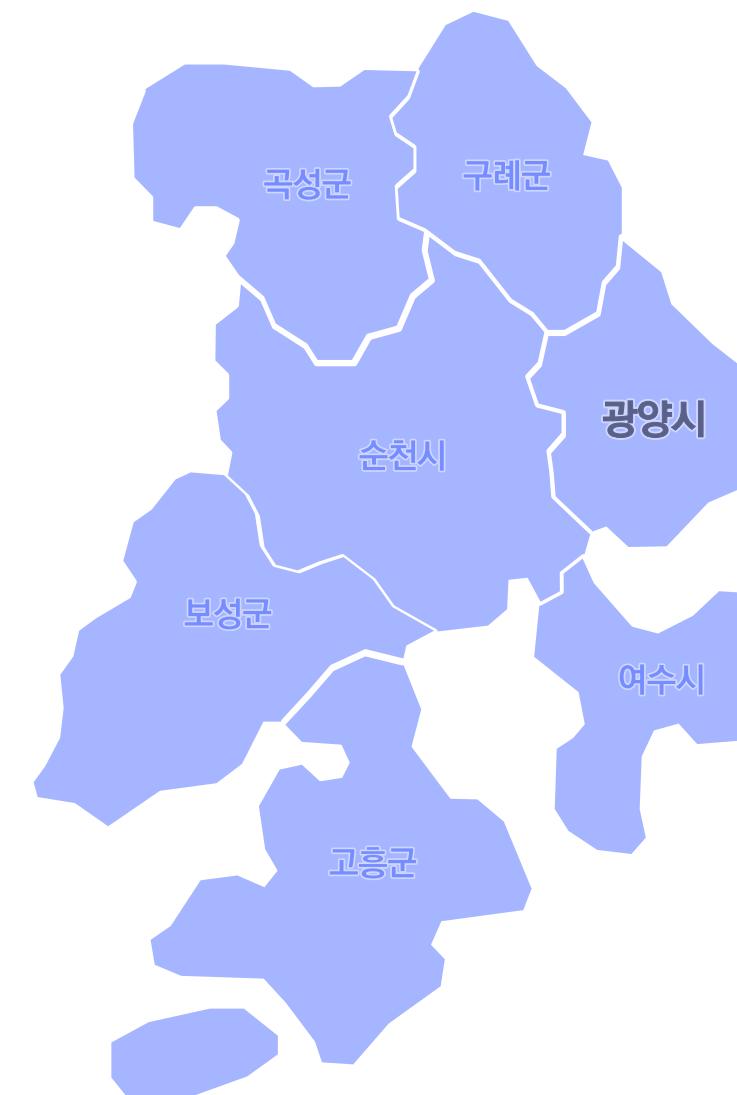
3. 세대당 인구(명)

4. 65세이상 고령자(명)

5. 인구밀도 ( $\text{명}/\text{km}^2$ )

6. 면적 ( $\text{km}^2$ )

7. 연도별 쓰레기 발생량

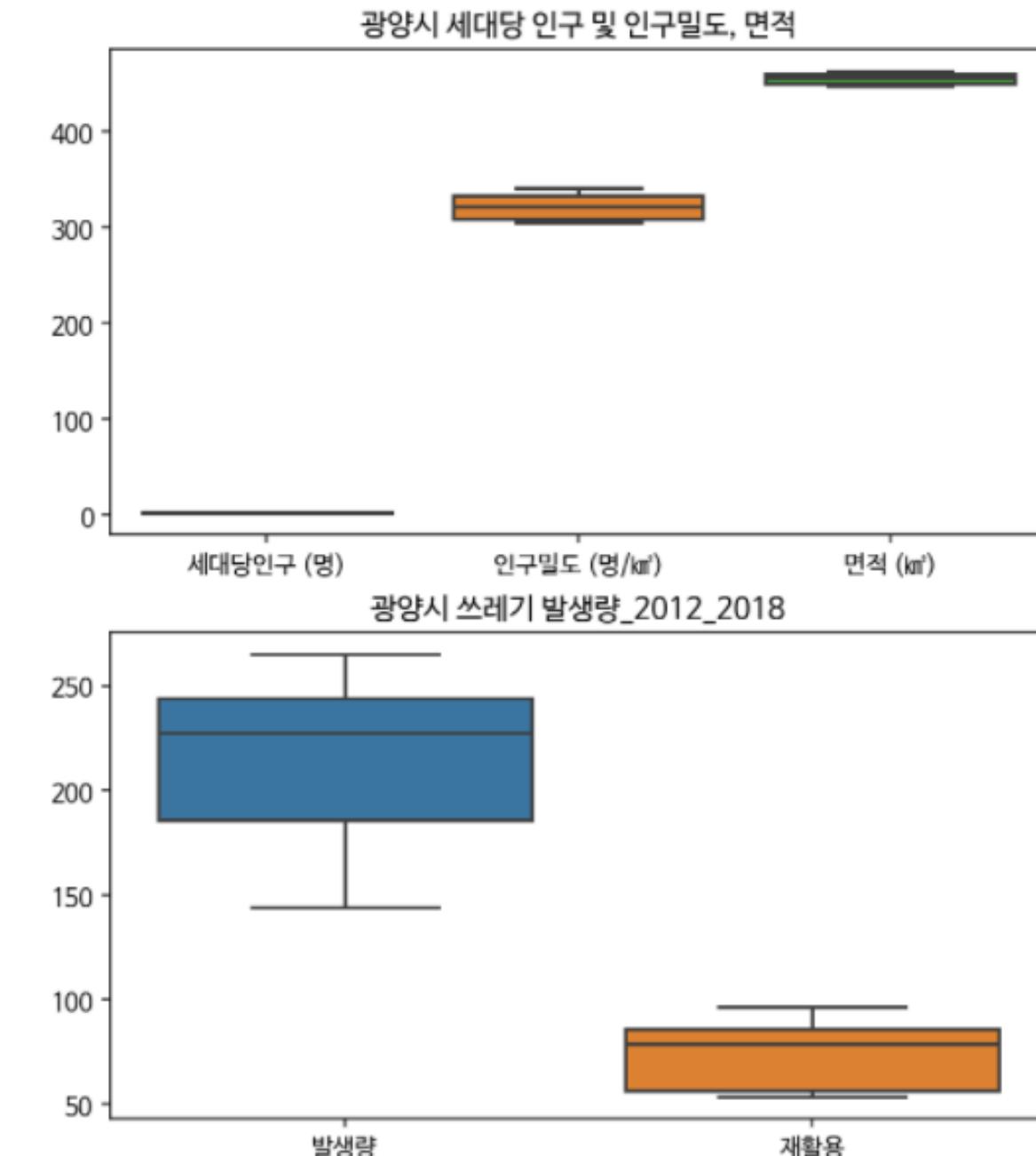
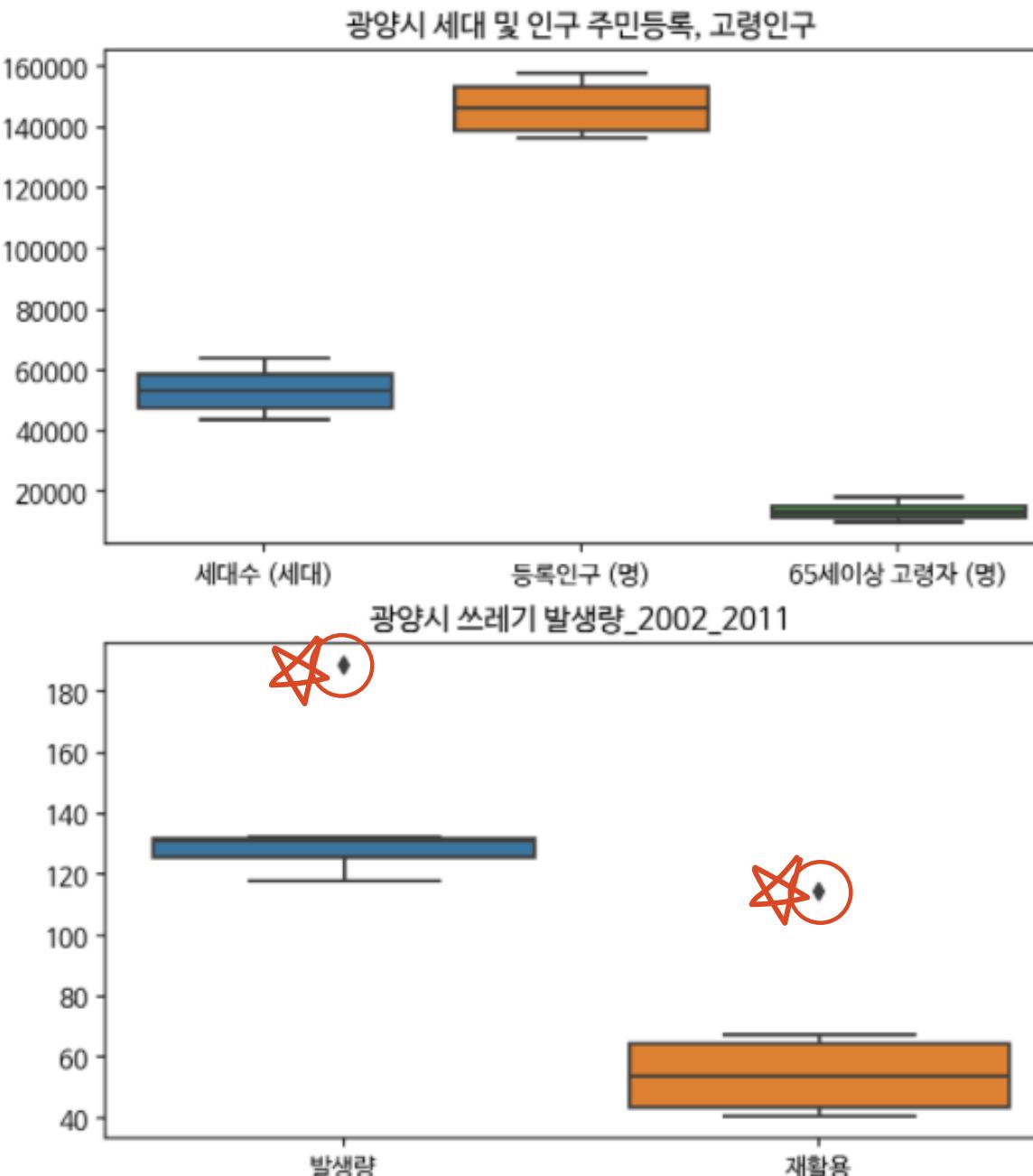


# 02 EDA / 전처리

01 데이터 불러오기

02 데이터 전처리

## 이상치 탐색 / 대체

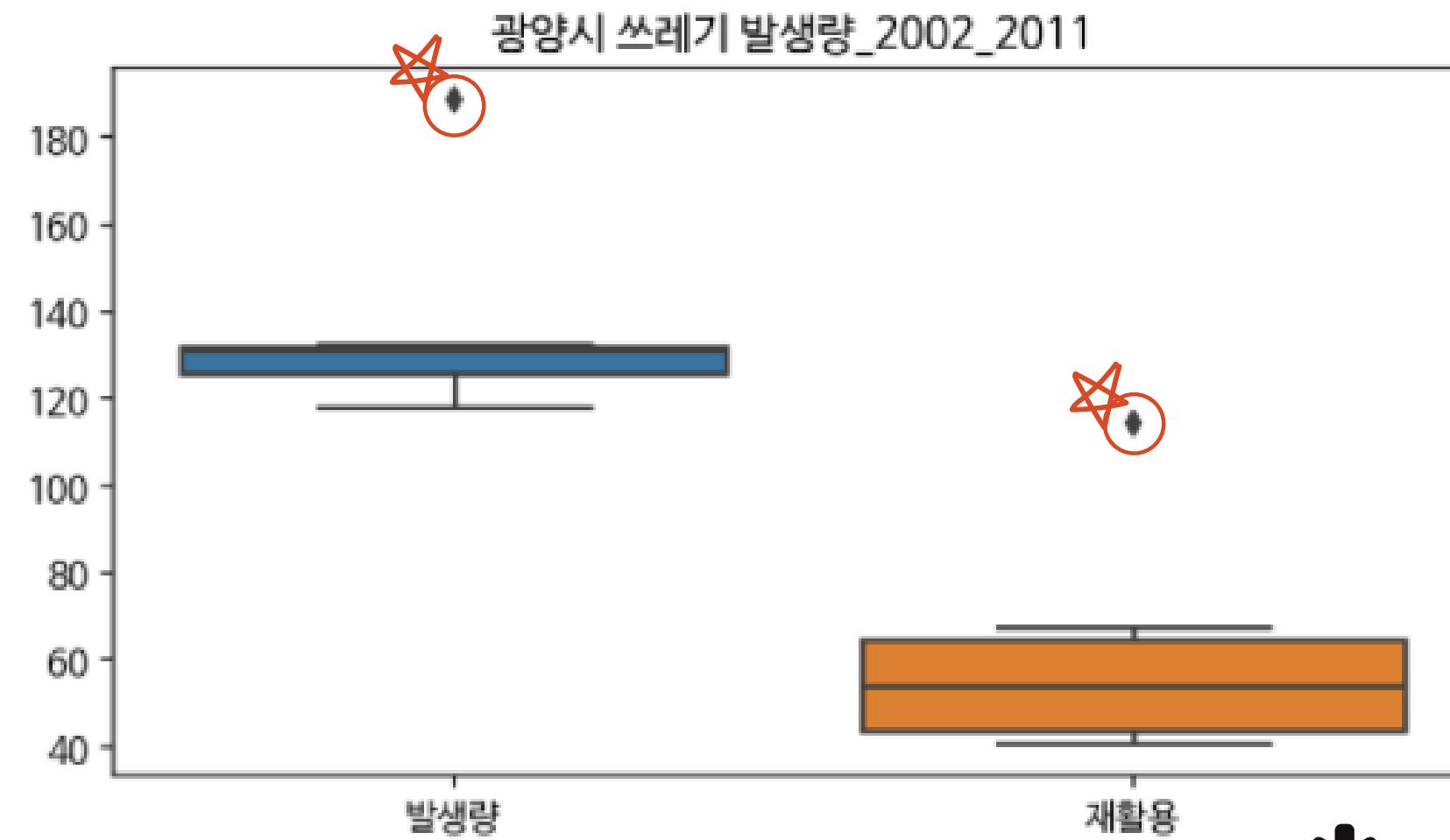


# 02 EDA / 전처리

01 데이터 불러오기

## 02 데이터 전처리

### 이상치 탐색 / 대체



seabom 라이브러리를 통한 boxplot 시각화로 이상치 발견.

기존 년도의 쓰레기 발생량 추이가 선형성을 띠기 때문에  
이상치를 보간법(interpolate)으로 대체.

# 02 EDA / 전처리

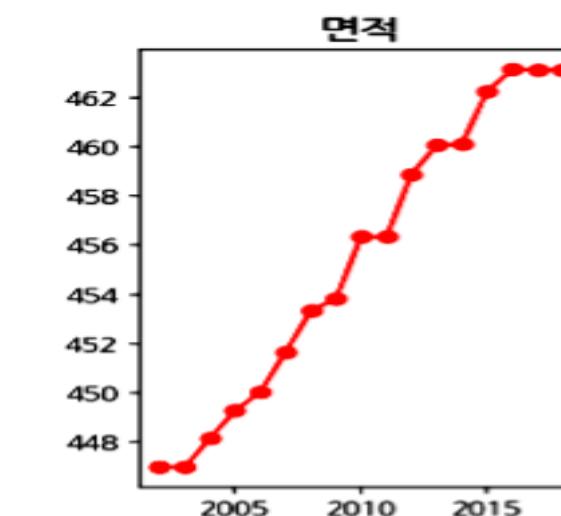
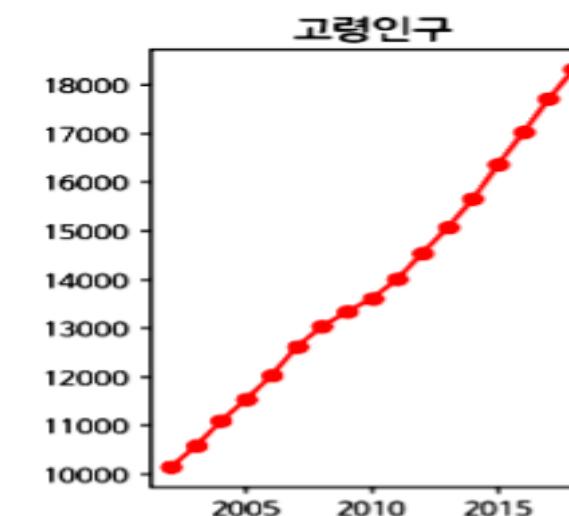
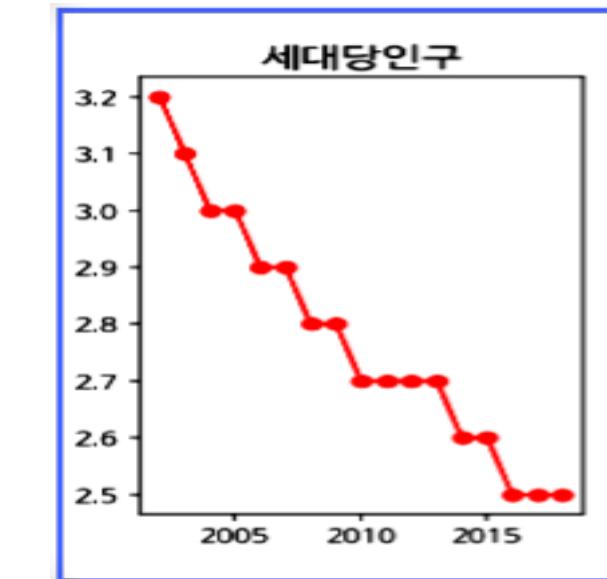
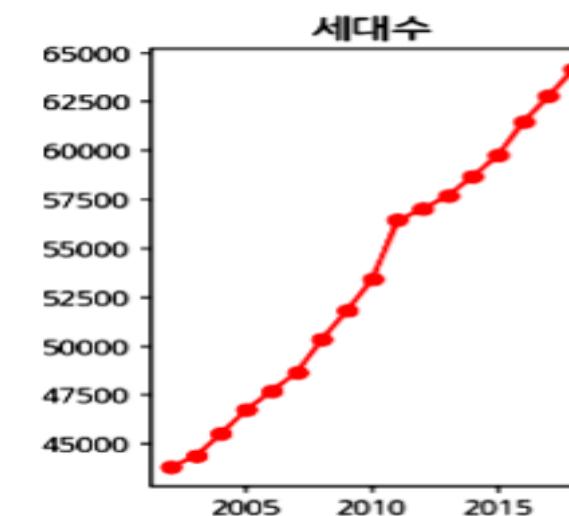
01 데이터 불러오기

## 02 데이터 전처리

# 데이터 경향 파악

각 특성별로 시각화하여 경향을 파악

시간흐름에 따라 대부분의 특성이 양의 상관관계



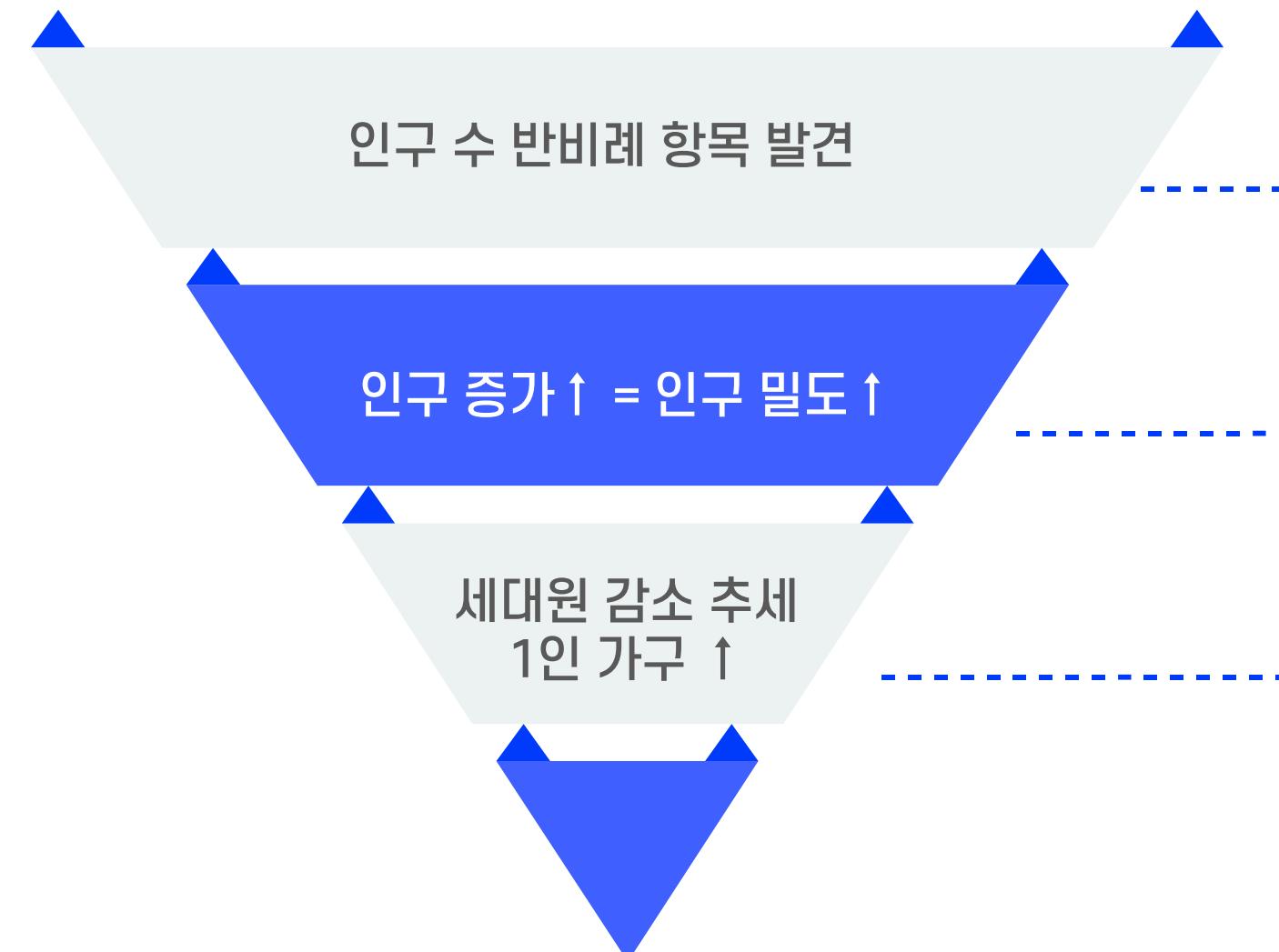
세대당 인구 특성 만  
음의 상관관계

# 02 EDA / 전처리

## 01 데이터 불러오기

## 02 데이터 전처리

# 데이터 경향 파악



◎ 세대 수, 고령 인구 수, 면적은 일정하게 증가  
세대 별 인구 수는 일정하게 감소

◎ 2011년 급격한 인구 증가가 있었고  
이는 인구밀도의 급격한 상승에도 영향을 주게 됨

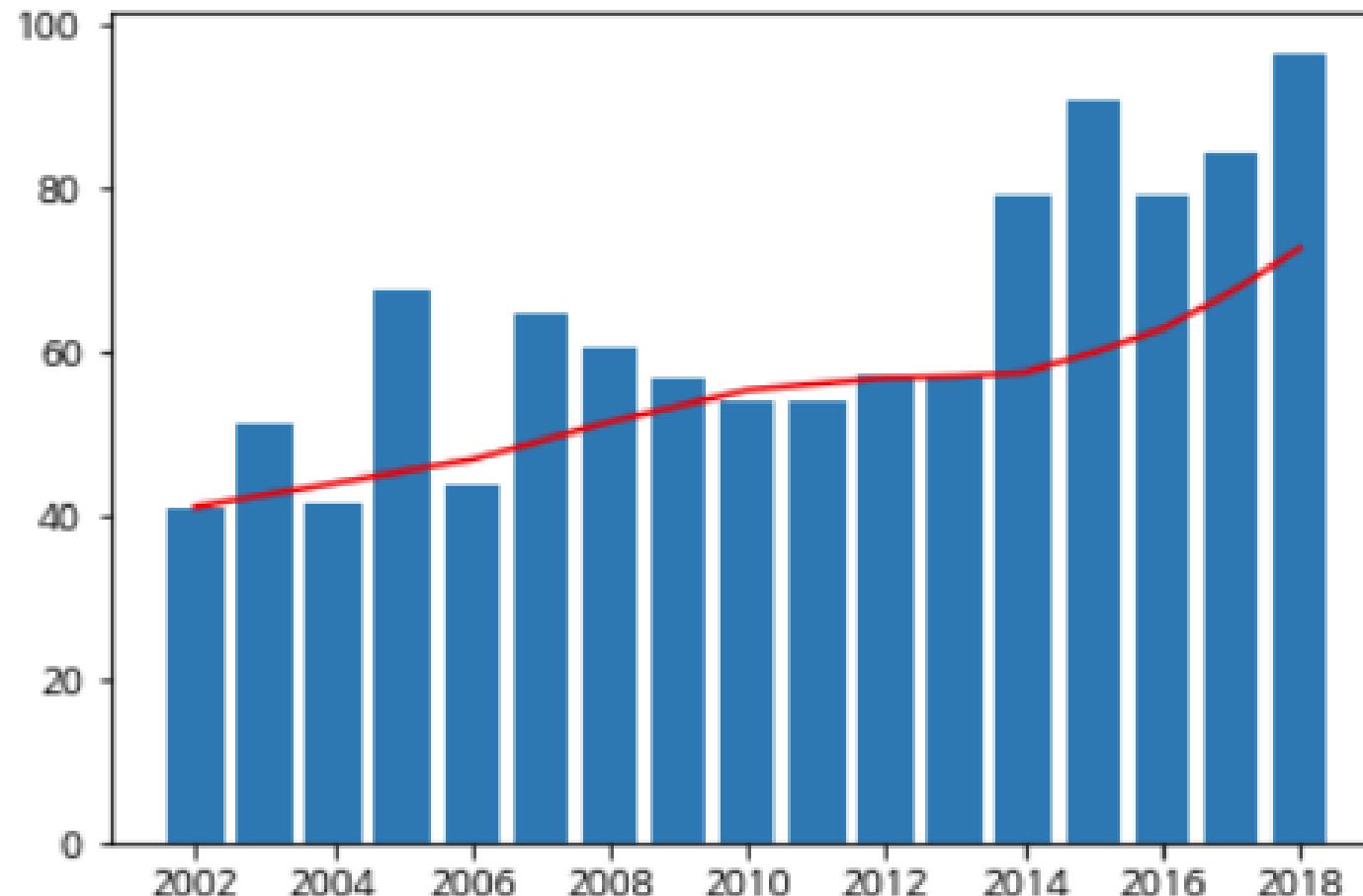
◎ 세대 별 인구 수가 지속적으로 감소하는 것으로 볼 때  
1인 가구 수는 증가하고  
3인, 4인 이상 가구 수는 지속적으로 감소

# 02 EDA / 전처리

## 데이터 경향 파악

01 데이터 불러오기

02 데이터 전처리



쓰레기 발생량의 연도별  
경향을 알아보기 위해  
이동평균선 산출

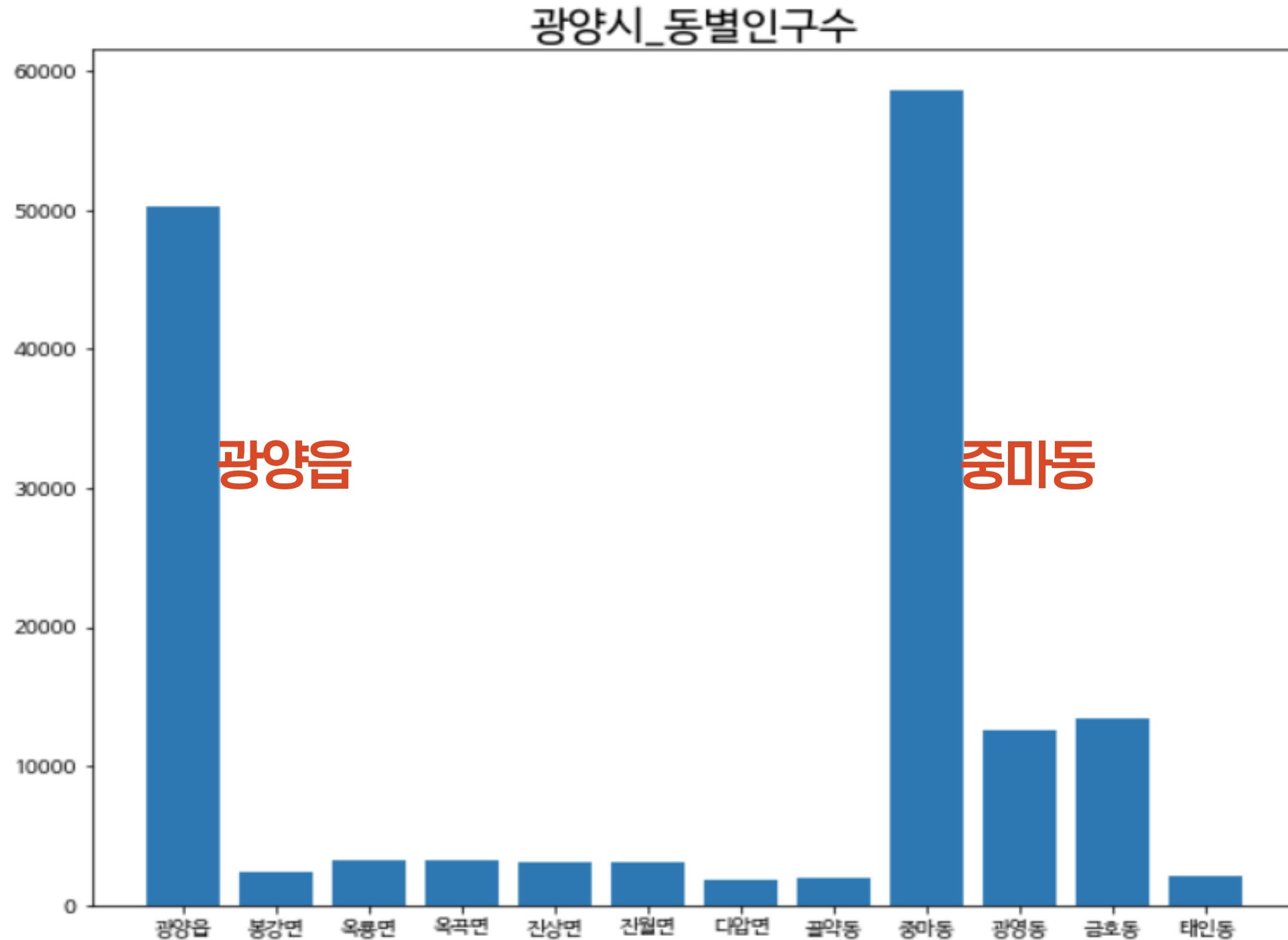
→ 증가 추세 확인

# 02 EDA / 전처리

## 인구분포 파악

01 데이터 불러오기

02 데이터 전처리

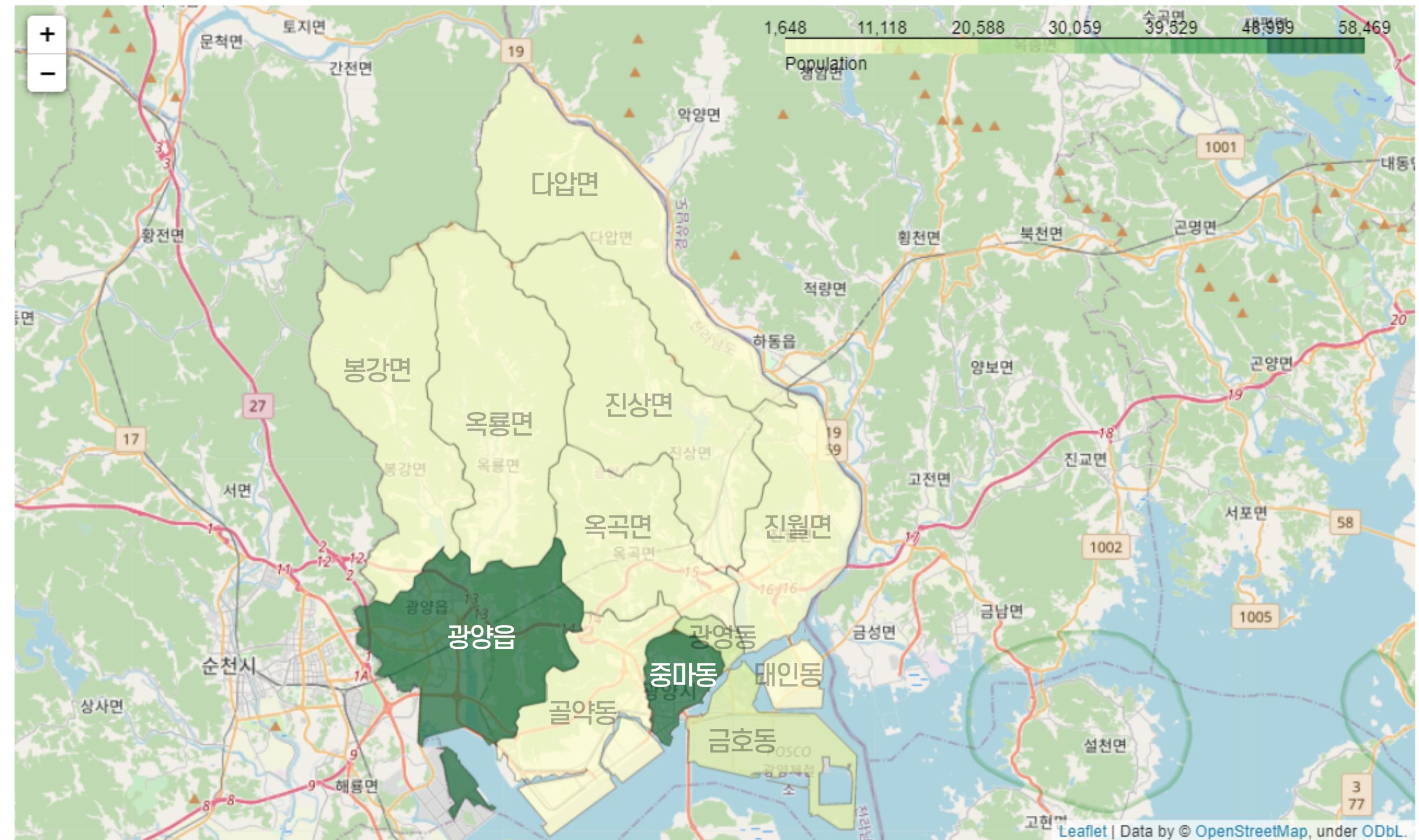


# 02 EDA / 전처리

# 인구 분포 파악

## 01 데이터 불러오기

## 02 데이터 전처리

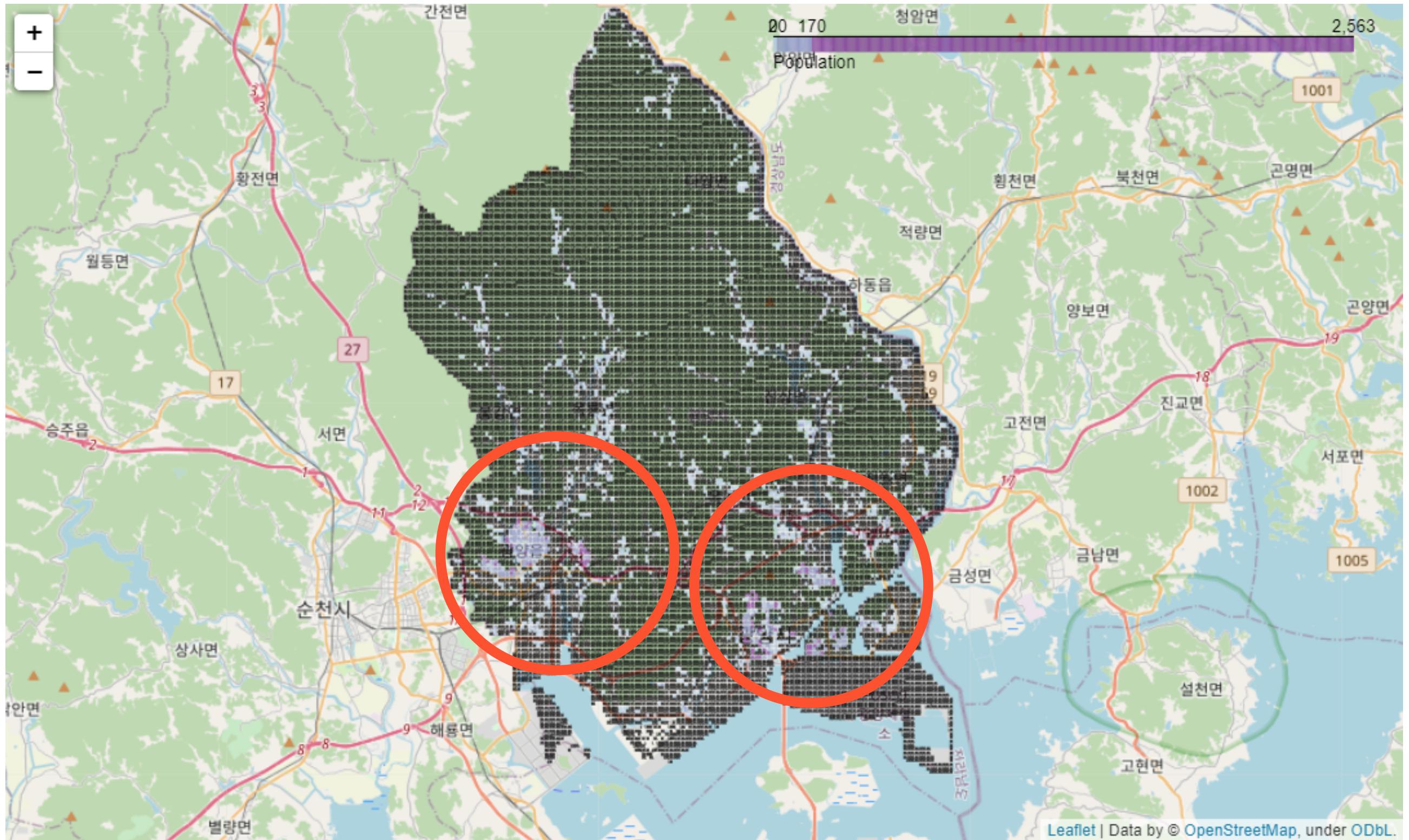


# 02 EDA / 전처리

01 데이터 불러오기

02 데이터 전처리

## 인구분포 파악(2)



## 02 EDA / 전처리

01 데이터 불러오기

### 02 데이터 전처리

## 인구 분포 경향

인구집중도

중마동 > 광양읍 > 금호동 > 광영동

인구분포는 위 4개의 행정구역 위주로 밀집되어 있으며 다른 지역에선 밀집되어 있지 않고 넓게 퍼져있는 경향

# 02 EDA / 전처리

01 데이터 불러오기

## 02 데이터 전처리

### 데이터 특성 중요도 분석

분석 목적

쓰레기 발생에 영향을 미치는 특성 중요도 파악

분석 과정

- MinMaxScaler로 데이터를 정규화한 후 특성 분석
- RandomForestRegressor 모델로 특성 중요도 파악 및 시각화
- 이동평균선 산출값이 경향성을 대표, 이동평균선 데이터로 학습 진행

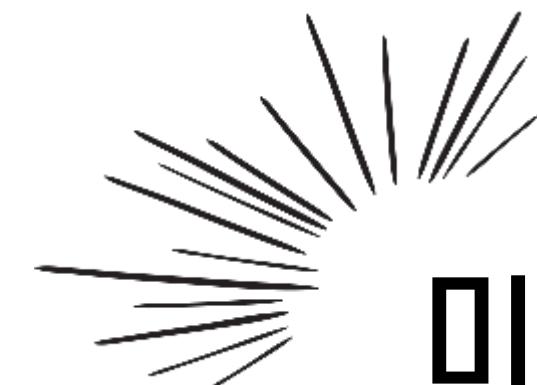
# 02 EDA / 전처리

01 데이터 불러오기

## 02 데이터 전처리

### 데이터 특성 중요도 분석

랜덤포레스트를 예측이 아닌  
특성 중요도 파악에만 사용한 이유 ??



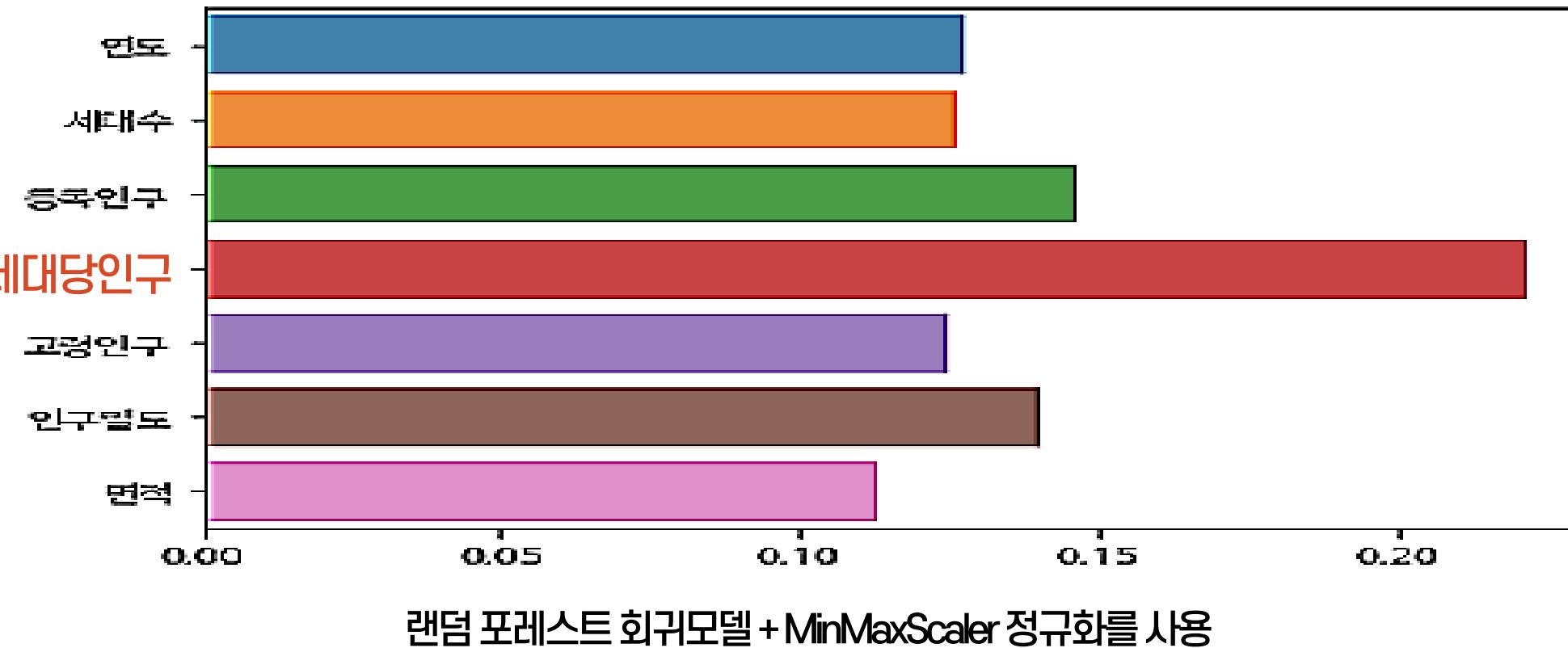
미래 예측에 입력할 데이터가  
학습에 사용했던 데이터 범위에서  
벗어나고 데이터 수가 **부족하기** 때문

# 02 EDA / 전처리

01 데이터 불러오기

## 02 데이터 전처리

### 인구 분포 경향



- 분석 결과
1. 특성중요도는 **세대당인구**에서 가장 높음.
  2. 1인 가구의 증가가 **쓰레기 발생량**에 큰 영향을 줌.
  3. 상대적으로 특성중요도가 낮은 **고령인구와 면적**을 제외

# 03

## 위치 선정을 위한 모델링

01. 목적 / 전처리
02. 필터링
03. 모델링 기법

# 03

위치 선정을  
위한 모델링

## 01 목적 / 전처리

02 필터링

03 모델링 기법

# 모델링 목적 / 사용 모델

## ※ 모델링 목적

1. **다중 선형 회귀 분석모델**을 행정구역별 미래 예측데이터에 적용하여  
**5년간 쓰레기 발생량 예측**
2. 작성한 모델을 행정구역별 데이터에 적용하여  
**지역별 쓰레기 발생량 예측** 및 시각화

## ※ 사용모델

- **다중 선형 회귀 분석 모델 (Linear Regression)** - 데이터 검증 목적
- **다항 회귀 모델 (Polynomial Regression)** - 데이터 예측 목적

# 03

위치 선정을  
위한 모델링

## 01 목적 / 전처리

02 필터링

03 모델링 기법

# 다중 선형 회귀 분석 모델 (LinearRegression)

- 다중선형회귀분석 모델(LinearRegration) 적용

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_i X_i$$

- 이동평균선 산출값이 경향성을 대표하므로 이동평균선 데이터로 학습 진행

- 각 항목의 값이 0이면 쓰레기 발생이 없으므로 `fit_intercept=0` 적용

- 각 계수들은 모두 쓰레기 증가 요인으로 `positive=true` 적용

- 과적합 방지를 위해 양상을 학습의 `bagging` 기법 적용

- 임의분할 교차검증을 이용해 train, test 진행, 스코어 검증

# 03

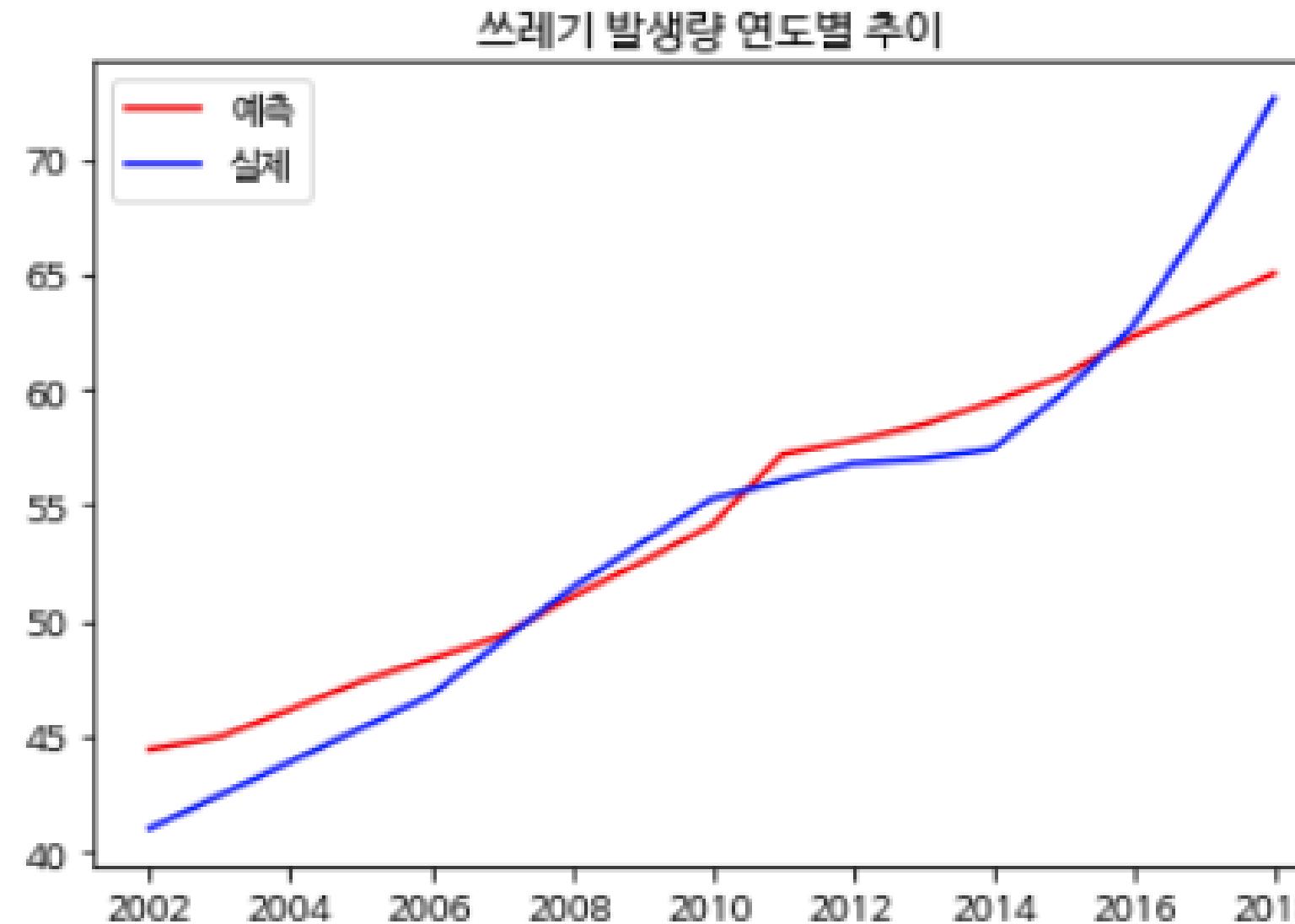
위치 선정을  
위한 모델링

## 01 목적 / 전처리

02 필터링

03 모델링 기법

# 다중 선형 회귀 분석 모델 (Linear Regression)



- 임의분할 교차검증을 진행하여 검증 신뢰도를 높였고, 정확도는 **91.8%** 도출
- 시각화 결과 회귀모델이 이동평균선의 경향성을 다소 잘 반영함

# 03

## 위치 선정을 위한 모델링

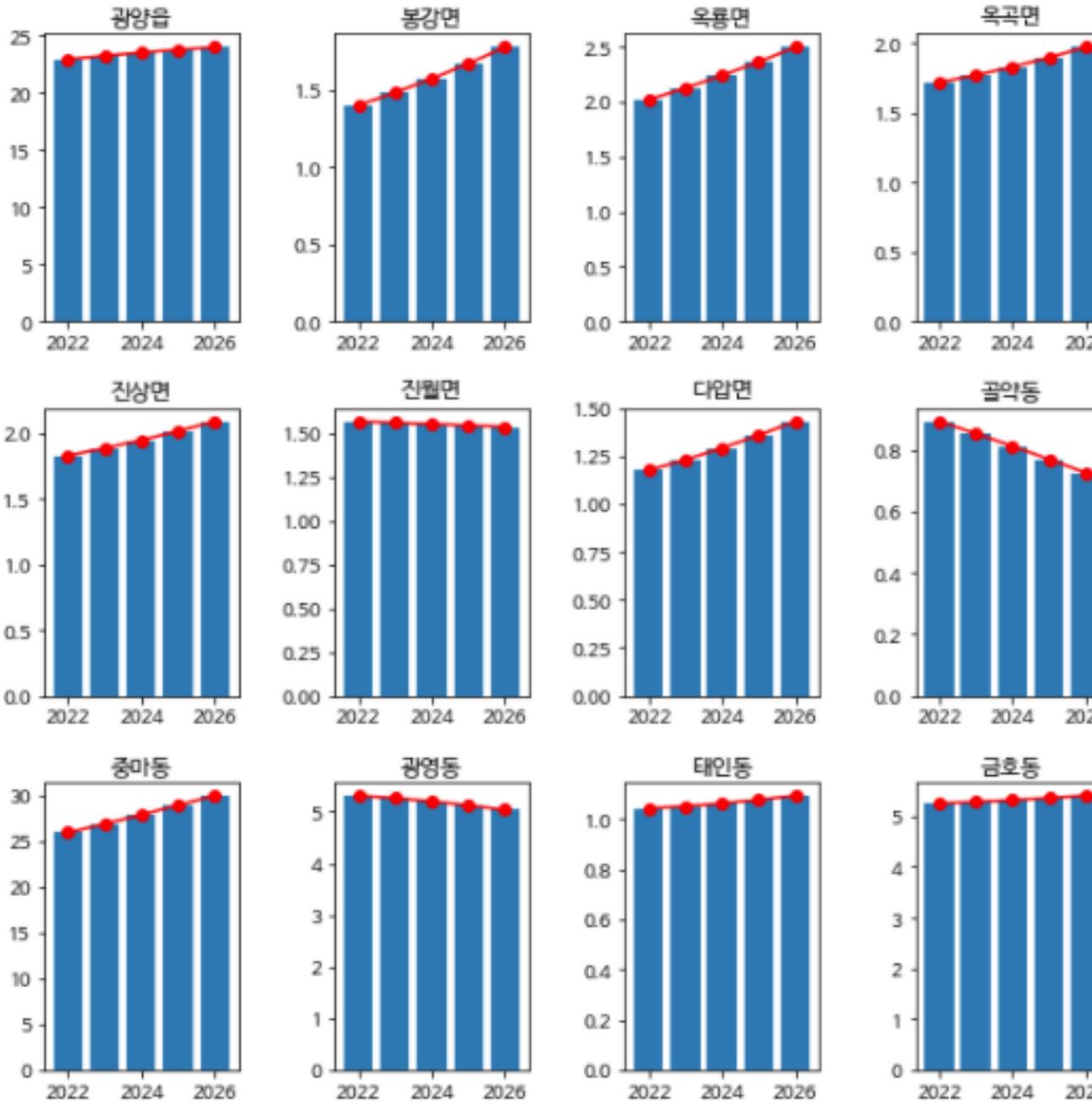
### 01 목적 / 전처리

### 02 필터링

### 03 모델링 기법

# 다항 회귀 모델 (Polynomial Regression)

동별 재활용 쓰레기 발생량 예측



- 연도에 따른 이차 회귀 모델을 이용, 행정동별 각 특성요소에 대한 미래 데이터 예측

- 미래 예측 데이터를 앞서 작성한 다중선형회귀분석 모델에 적용

- 미래 5년 (2022-2026)의 행정구역별 쓰레기 발생량 예측 모델 완성

# 03 위치 선정을 위한 모델링

## 01 목적 / 전처리

## 02 필터링

## 03 모델링 기법

# 다항 회귀 모델 (Polynomial Regression)

행정동	동별 5년간 쓰레기 발생 예측량 평균
0 중마동	27.887943
1 광양읍	23.460608
2 금호동	5.326887
3 광영동	5.188990
4 옥룡면	2.246448
5 진상면	1.954469
6 옥곡면	1.836584
7 봉강면	1.579559
8 진월면	1.552918
9 다압면	1.296746
10 태인동	1.065793
11 골약동	0.809947

- 행정구역별 미래 쓰레기 발생량을 내림차순으로 조회해본 결과, **중마동, 광양읍, 금호동, 광영동** 이상 4개의 읍면동이 전체 12개 읍면동의 83.3%를 차지함
- 행정구역별 5년간 쓰레기 발생 예측량 평균으로 상위 4개의 행정동을 추출
- 금호동은 12개 고정 남은 38개는 쓰레기 발생량에 따라 분배

# 03

위치 선정을  
위한 모델링

## 01 목적 / 전처리

02 필터링

03 모델링 기법

## 모델링 예측 결과

- 광양읍, 중마동과 같이 인구수가 많고 인구밀도가 높은 지역은 쓰레기 발생량이 증가하는 경향을 보임.
- 인구수가 적은 지역은 큰 변화가 없거나 쓰레기 발생량이 감소하는 경향 확인.
- 재활용 쓰레기 발생량이 가장 많은 4개 행정동은 광양읍, 중마동, 금호동, 광영동으로 예측
- 재활용 쓰레기 발생량 비례하여 수거기 개수지정  
광양읍 16, 중마동 19, 금호동 12개 광영동 3개의 수거기를 배치하기로 결정

# 03

위치 선정을  
위한 모델링

## 01 목적 / 전처리

02 필터링

03 모델링 기법

# 데이터 전처리 및 세부 분석 과정

- ◎ 건물 정보 중 공동주택만 필터링
- ◎ 필터링한 데이터 행정구역별로 구분
- ◎ 각 건물의 중심 좌표 시각화
- ◎ 건물 표제부 정보 시각화
- ◎ 행정구역별, 공동주택 및 세대수 상위20% 필터링

## [사용된 데이터]

4.광양시\_건축물\_표제부.csv

7.광양시\_도로명주소(건물).geojson

20.광양시\_행정경계(읍면동).geojson

# 03

위치 선정을  
위한 모델링

01 목적 / 전처리

## 02 필터링

03. 모델링 기법

# 공동 주택 필터링

최종적으로 선정되어야 할 공동 주택을 위해  
기존 데이터의 필터링이 필요함

(※ 광영동의 다세대 주택은 군집화에서  
이상치로 작용할 가능성이 높아 제외)



광양시 전체 건물에서

광양시 전체의 공동주택을  
행정구역 별로 필터링함

# 03

위치 선정을  
위한 모델링

01 목적 / 전처리

02 필터링

03. 모델링 기법

## 공동 주택 필터링

인구수와 재활용쓰레기 발생량이 많은 **행정 구역**에

공동주택이 많이 분포해있음을 확인

광양읍	중마동	금호동	광영동
-----	-----	-----	-----



**세대수를 기준으로  
상위20% 필터링**



광양읍	중마동	금호동	광영동
-----	-----	-----	-----

# 03 위치 선정을 위한 모델링

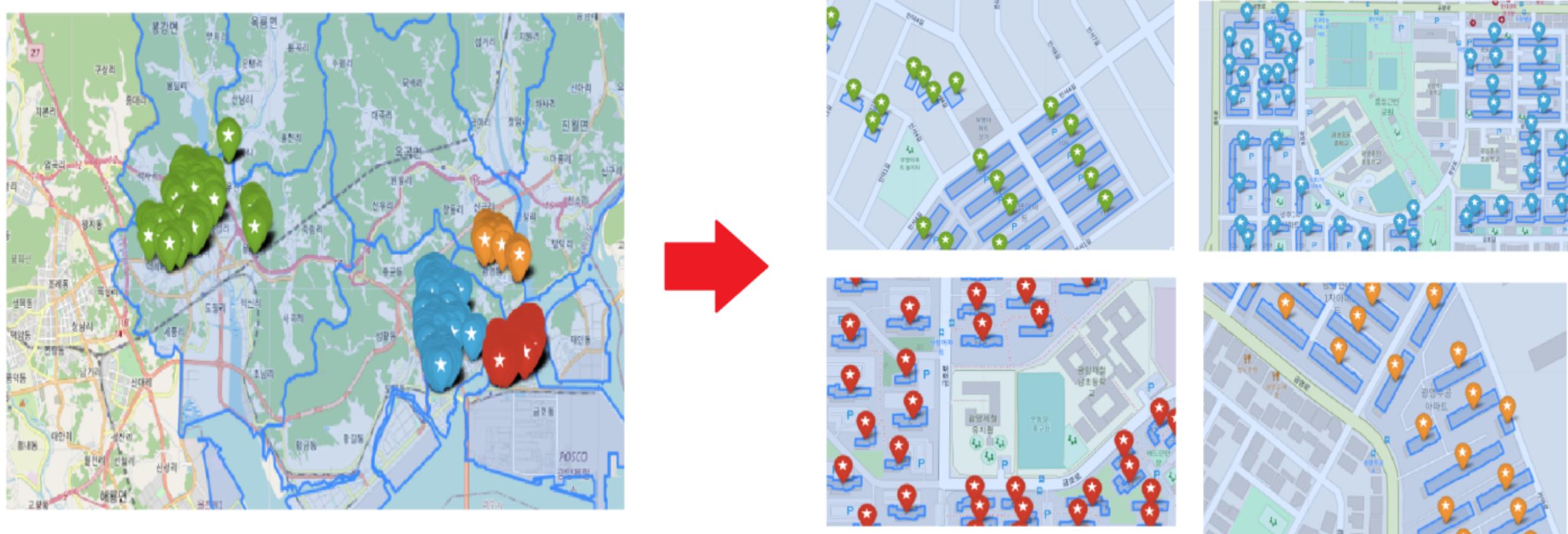
01 목적 / 전처리

02 필터링

03.모델링 기법

## 행정구역별 공동 주택 중심좌표

※ 수거기 설치 가능 장소를 식별하기 위해  
행정구역별로 세분화 후 각 건물의 중심좌표 시각화



# 03

위치 선정을  
위한 모델링

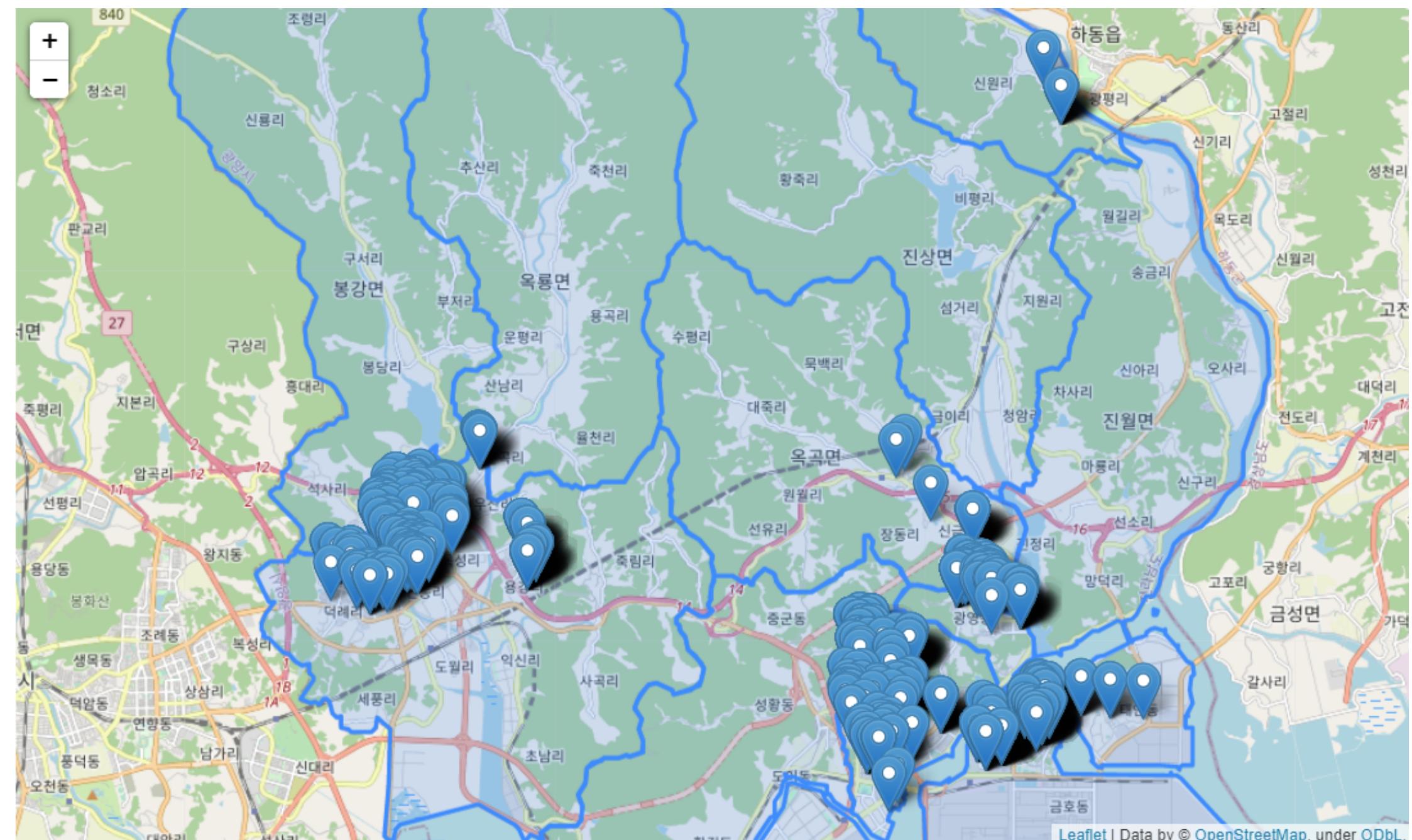
01 목적 / 전처리

02 필터링

03. 모델링 기법

## 건물 표제부 정보

※ 건물 표제부 정보 시각화



# 03 위치 선정을 위한 모델링

01 목적 / 전처리

02 필터링

03.모델링 기법

## 모델링 기법

### [사용된 모델]

- ◎ KMeans
- ◎ MeanShift
- ◎ DBSCAN
- ◎ OPTICS
- ◎ SpectralClustering
- ◎ AgglomerativeClustering

[모델 스코어 평가]

silhouette\_score ,calinski\_harabasz\_score 이용

04

## 군집화

- 01. 성능 테스트
- 02. 모델 튜닝

# 04 군집화

## 01 성능 테스트

## 02 모델 튜닝

# 군집화 성능 테스트

클러스터의 중심좌표를 산출하기 위해

비지도 학습 **군집화** 진행하였고 군집화의 성능을 관찰함

인구수가 가장 많은 **중마동**을 테스트 대상으로 선정

성능평가관찰을 목표로 하기 때문에 가중치 없이

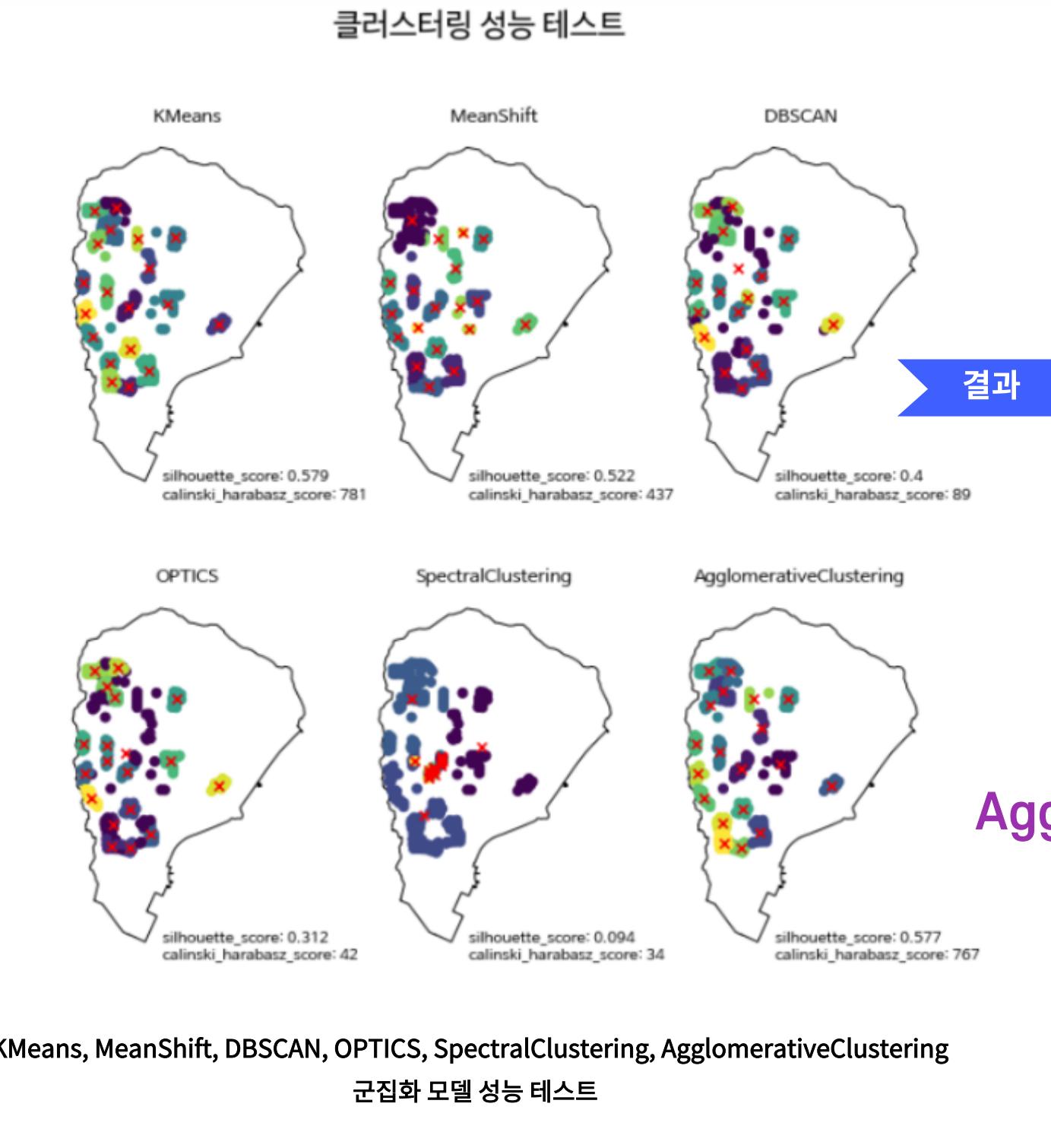
공동주택 중심좌표만 가지고 클러스터링 진행

# 04 군집화

## 01 성능 테스트

## 02 모델 튜닝

# 군집화 성능 테스트 결과



높은 성능	낮은 성능
KMeans	DBSCAN
MeanShift	OPTICS
AgglomerativeClustering	SpectralClustering

KMeans,  
MeanShift,  
AgglomerativeClustering → 비교적 좋은 군집화 성능

DBSCAN,  
OPTICS,  
SpectralClustering → 낮은 성능점수로 부적합 하다고 판단

# 04 군집화

## 01 성능 테스트

## 02 모델 튜닝

### 군집화 성능 테스트 결과

**silhouette\_score** : 1위: KMeans 0.579  
2위: AgglomerativeClustering 0.577

\*실루엣 지표는 각 군집 간의 거리가  
얼마나 효율적으로 분리되어 있는지 평가할 수 있음

**calinski\_harabasz\_score** : 1위: KMeans 781  
2위: AgglomerativeClustering 767



평가 지표를 기준으로 하여  
**[KMeans]**를 최종 모델로 선정함

# 04 군집화

## 01 성능 테스트

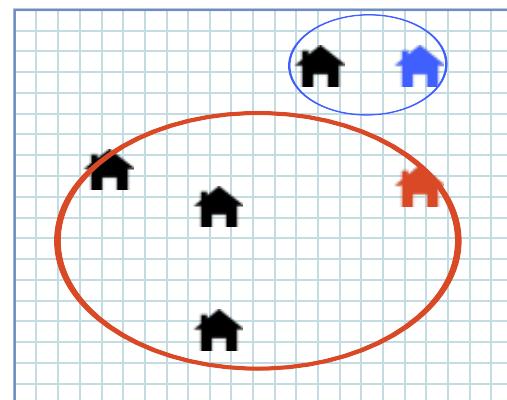
## 02 모델 튜닝

# KMeans 모델 원리

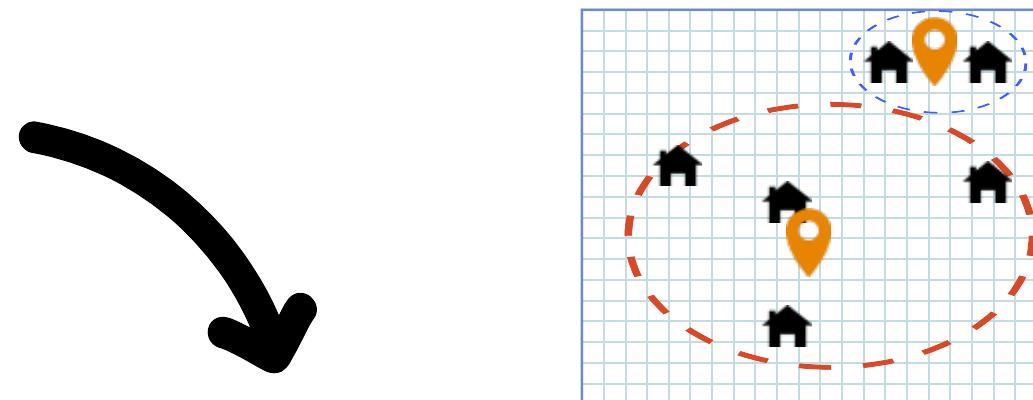
행정구역 별 공동주택의 중심 좌표 산출을 위해 군집화 수행함



랜덤하게 중심값 지정



새로운 중심 선택



중심변화가 없다면 (범위 확인) 정지  
변화가 있다면 과정 반복

K개의 중심값과 각 데이터간의  
거리 측정 및 할당



# 04 군집화

01 성능 테스트

## 02 모델 튜닝

### KMeans 모델 튜닝

[행정 구역 별로 지정된 수거기 개수  
= 클러스터 개수]



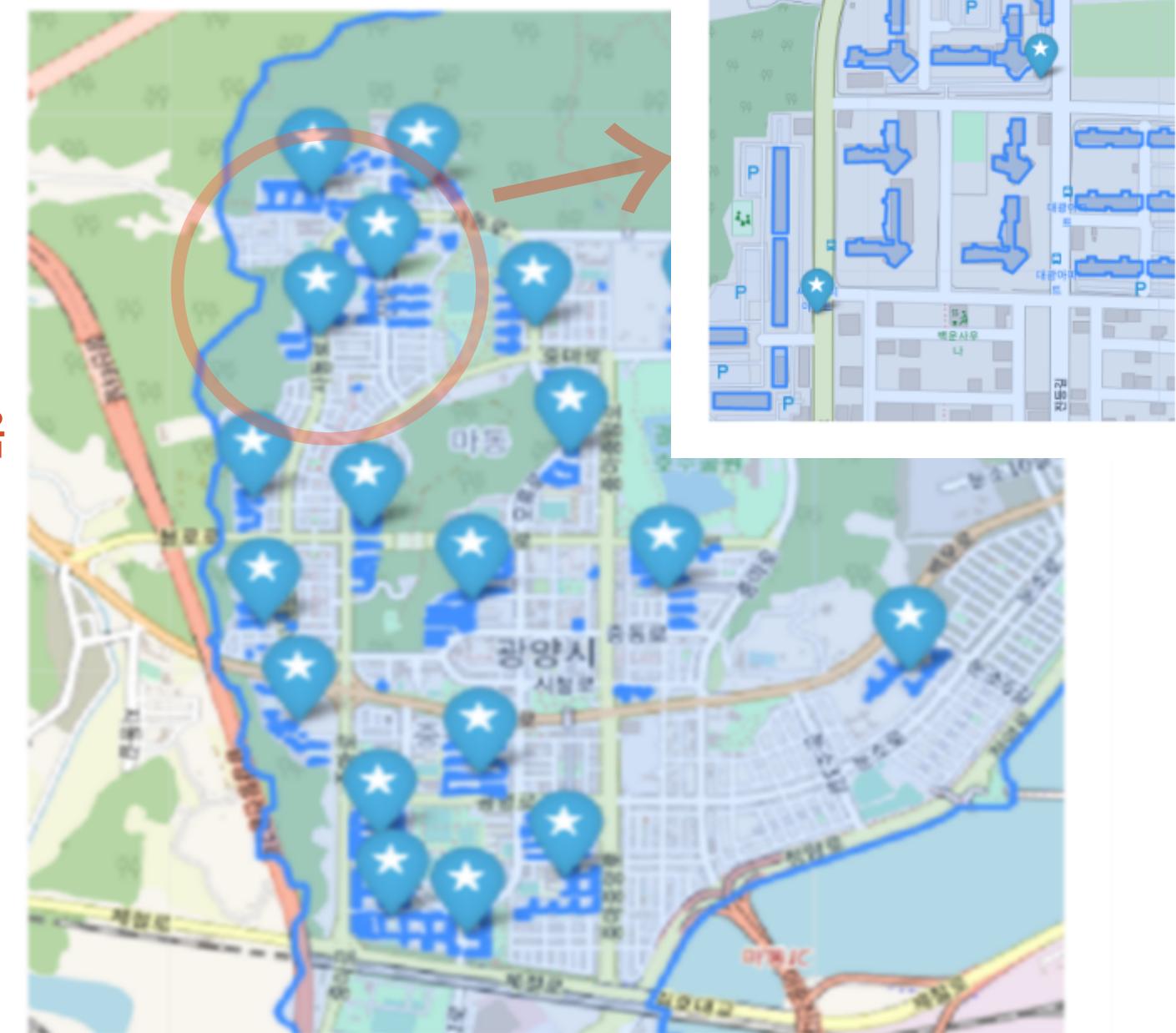
[오류]

클러스터 중심의 좌표가  
공동주택 내에 위치하지 않는 경우가 있음



[개선]

KMeans 모델을 **튜닝**하여  
클러스터 중심과 가까운 공동주택단지를  
수거기 위치좌표로 선정하고  
행정구역 별로 군집화 적용



# 04 군집화

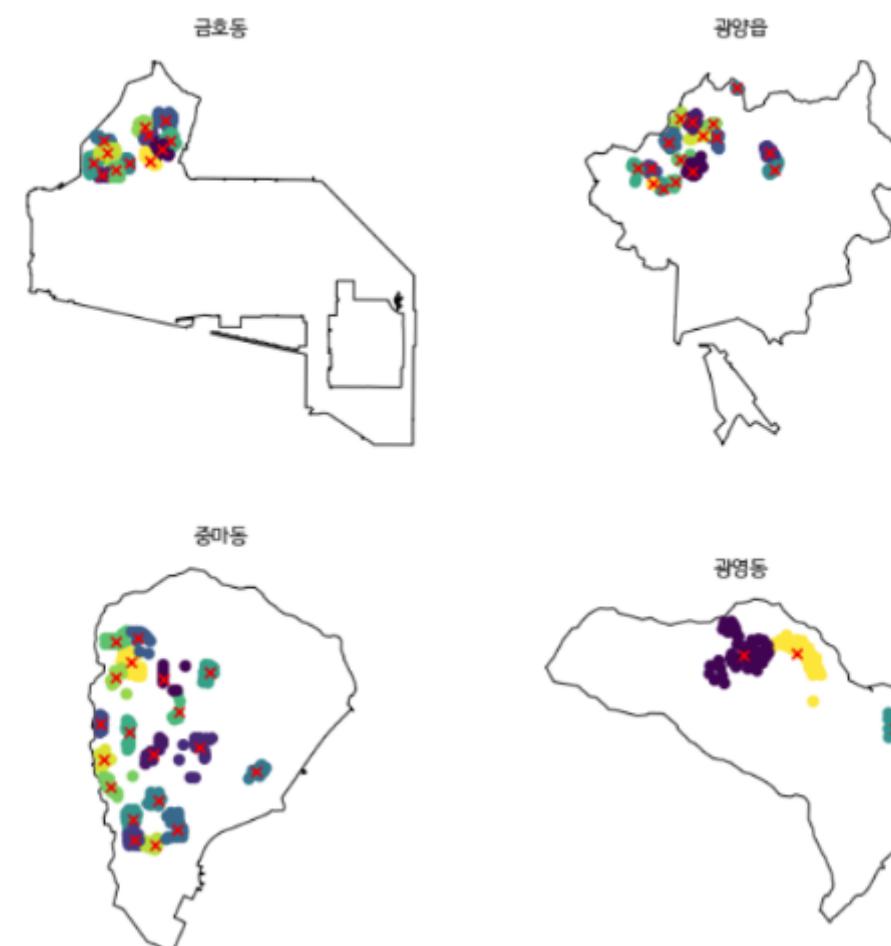
## 개선된 모델 사용

### 01 성능 테스트

### 02 모델 튜닝

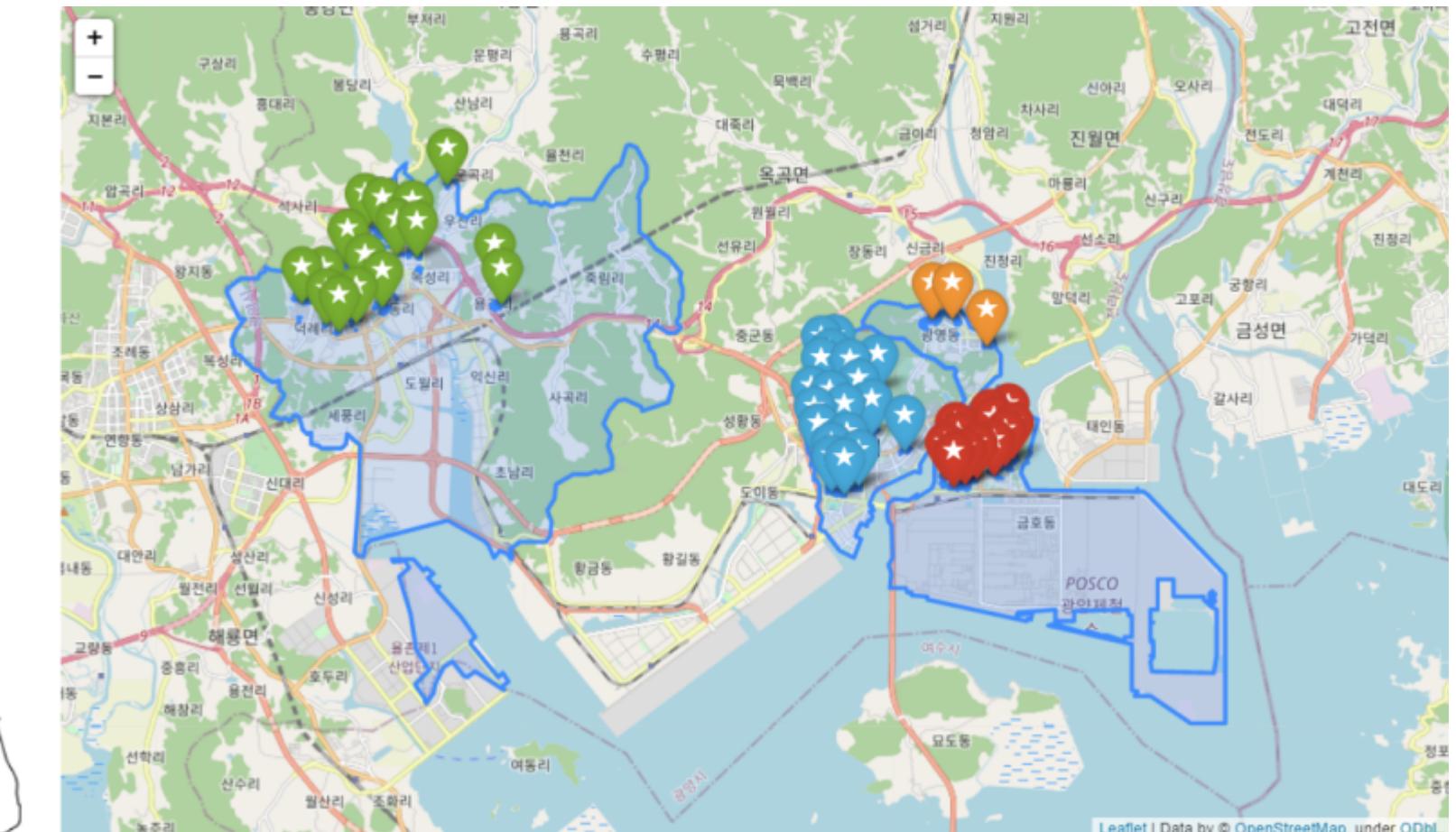
#### 클러스터링 결과 행정구역별 수거기 최종 위치 선정

동별 건물좌표 클러스터링



행정구역 별 건물 좌표

클러스터링 결과



광양읍 16개

중마동 19개

금호동 12개

광영동 3개

# 04 군집화

01 성능 테스트

## 02 모델 튜닝

# 개선된 모델 사용

## 장려금 수령에 관한 내용을 가중치 부여

한편, 시는 쓰레기 감량 및 자원순환 이용 촉진을 위해 재활용품 분리배출 지원 장려금을 지급하고 있다.

인공지능 캔/페트병 수거기를 광양읍, 중마동, 광영동 주민센터와 금호동 백운쇼핑센터에 설치해 이용자에게 장려금을 지급하며, 플라스틱, 페트병, 형광등 등 생활폐기물과 영농 폐비닐, 농약병 등 영농폐기물 분리 배출 시 kg당 30~1,000원의 장려금을 지급하고 있다.

## 제공된 파일 중

2.광양시\_재활용품분리수거\_장려금지급내역.csv에서  
장려금을 수령한 지역 중 상위 20%를 선정하여 가중치를 부여함

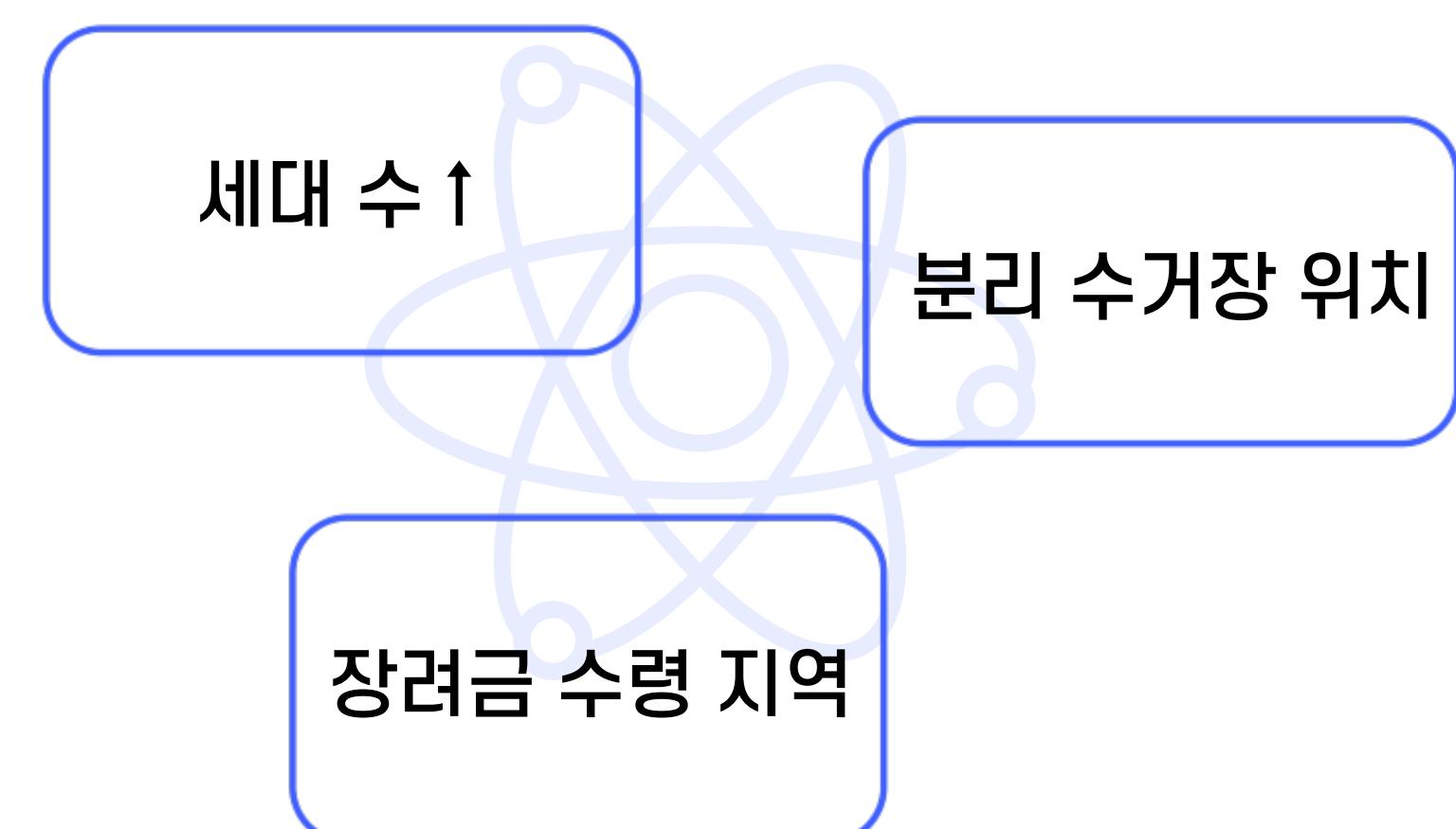
# 04 군집화

01 성능 테스트

02 모델 튜닝

## 가중치 부여

각 행정구역의 공동주택 중심좌표에  
위치선정의 의미를 부여하는 **가중치 부여**



최적의 위치

# 04 군집화

01 성능 테스트

02 모델 튜닝

## 가중치 부여

세대 수 ↑

각 공동 주택별 세대 수는 쓰레기 발생량과  
밀접한 관계가 있다고 판단하여 가장 큰 가중치

기존 공동 주택 내에 위치한 분리수거장은 주민들의  
동선을 고려하여 합리적으로 설치되었다고 판단

분리 수거장 위치

장려금 수령 지역

장려금 우수 지급 공동주택의 경우  
자동수거기 설치에 따른 효과가 커질 것으로 판단

# 05

## 결론

01. 공개 데이터
02. 비공개 데이터

# 05 결론

## 01 공개 데이터

02 비공개 데이터

# 군집화 대상 좌표 리스트 업데이트

세대수 \* 5



건물위치 + 5\*(세대 수 상위 20% 공동 주택)

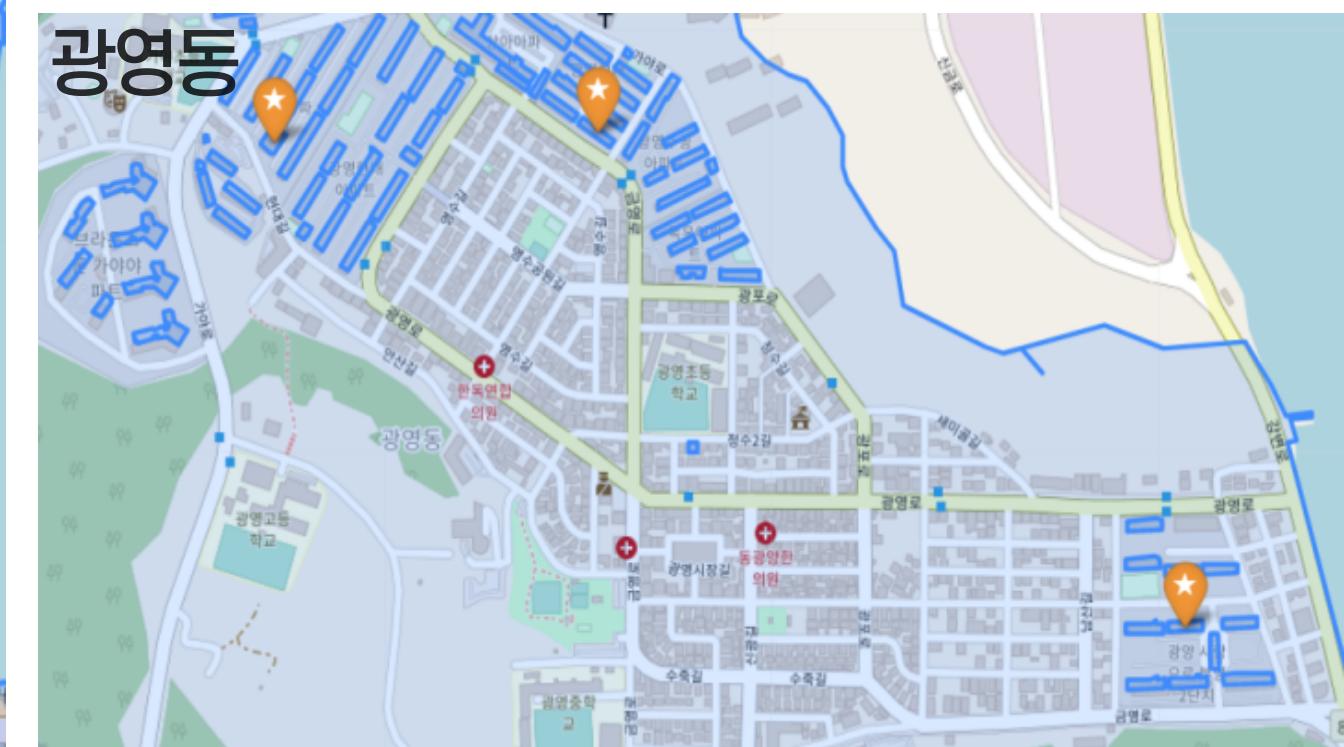
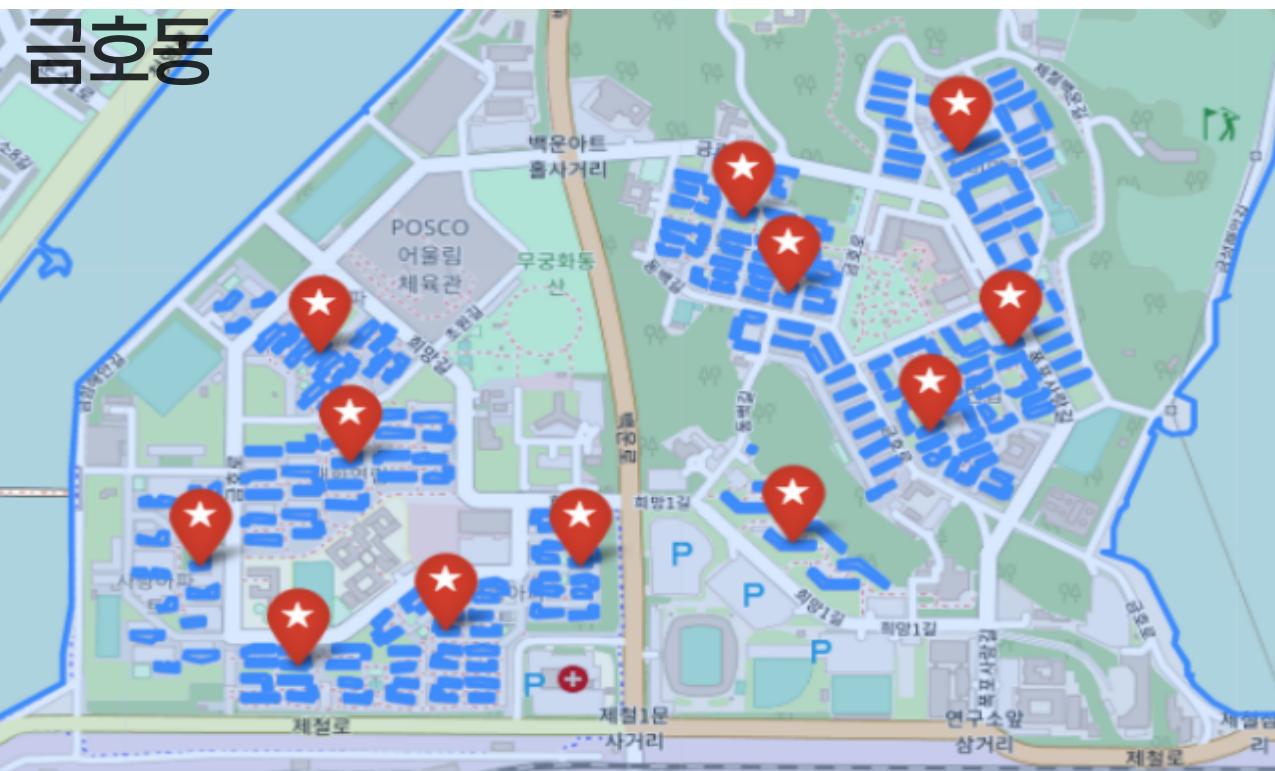
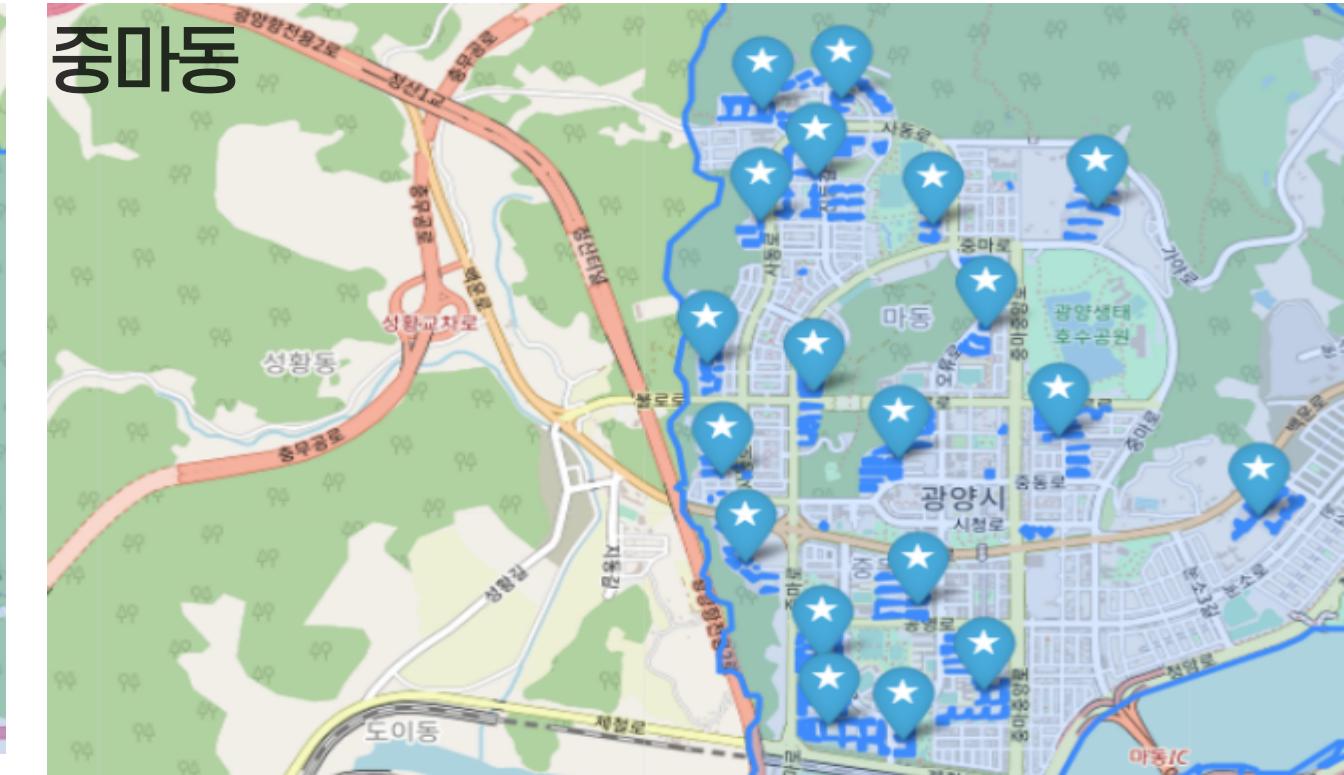
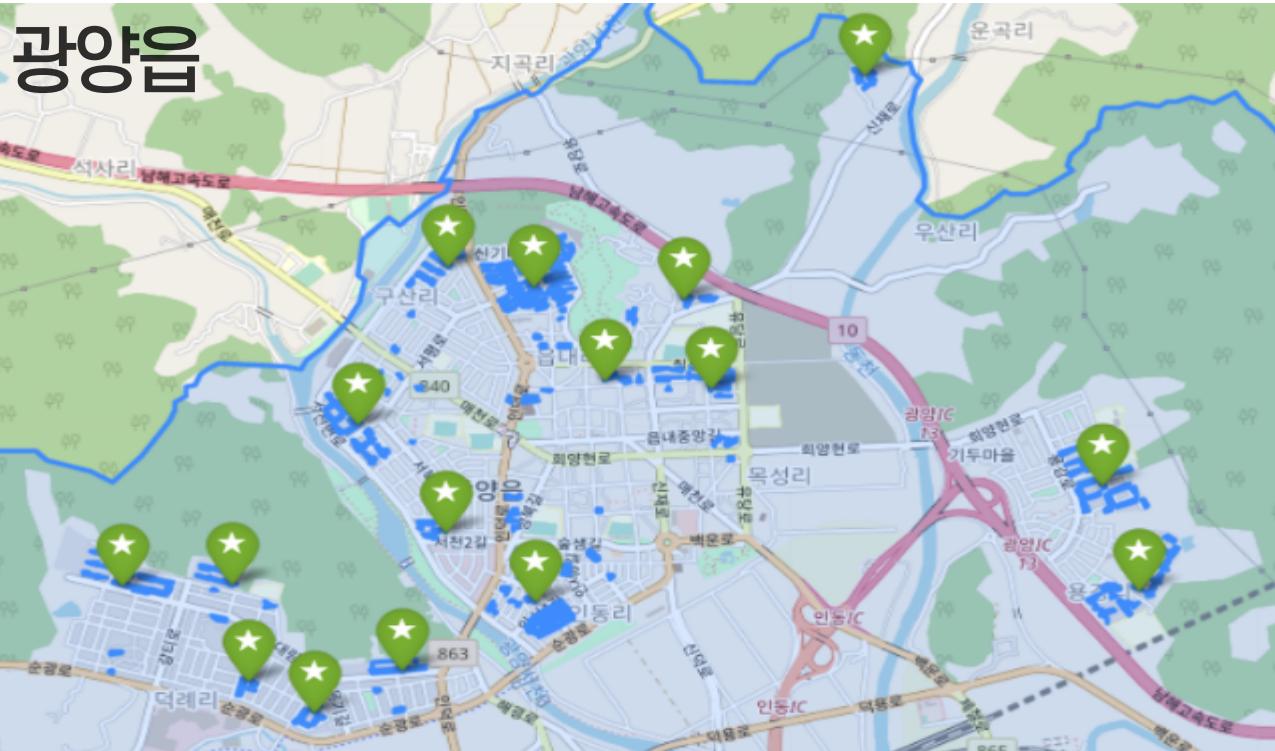
→ 공개 데이터로 가중치를 매길 요건은 인구 수에 국한

# 05 결론

## 01 공개 데이터

## 02 비공개 데이터

# 공개 데이터 좌표

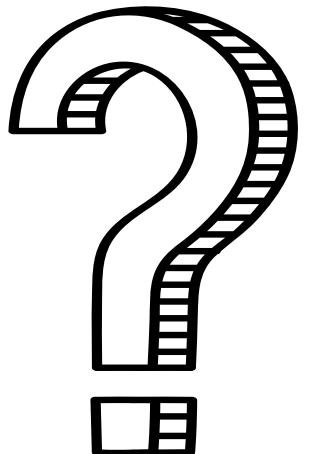


# 05 결론

01 공개 데이터

02 비공개 데이터

## 비공개 데이터 활용



COMPAS Notebook 외 다른 IDE로 비공개 데이터에 접근이 어렵기 때문에 비공개 데이터를 활용한 **가중치 적용** 결과를 따로 분석



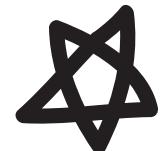
가중치 적용 전 가장 중요한 고려사항은  
**군집화 대상리스트 내에 같은 좌표를 여러개 겹쳐서 추가하는 것 !!**

# 05 결론

01 공개 데이터

02 비공개 데이터

## 비공개 데이터 활용



★ 가중치 적용 전 가장 중요한 고려사항은  
군집화 대상 리스트 내에 같은 좌표를  
여러개 겹쳐서 추가하는 것 !!



가중치가 높은 좌표에 클러스터링 중심이 가까워질 가능성이 높음

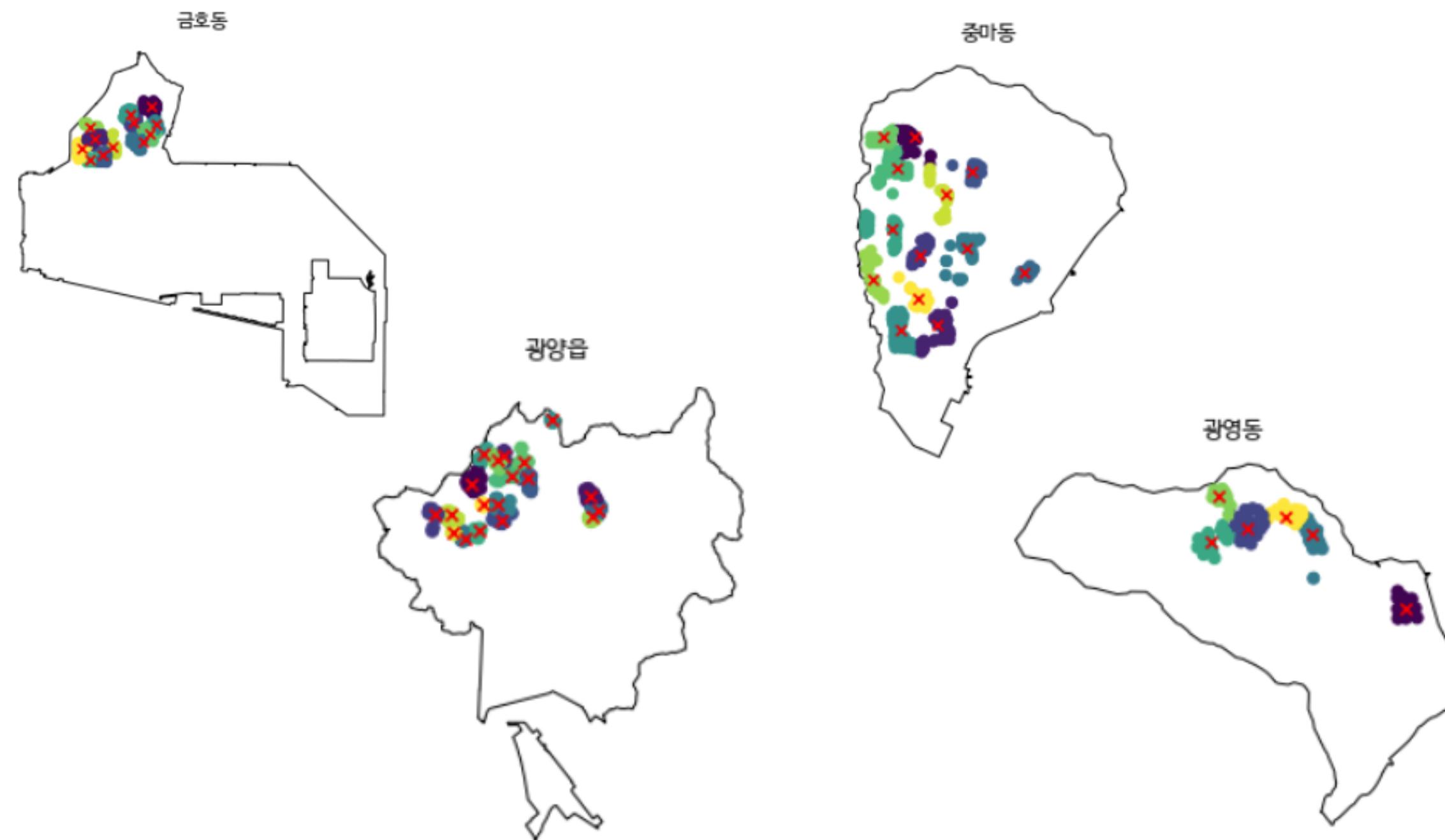
# 05 결론

01 공개 데이터

02 비공개 데이터

## 공동주택 분리수거장 현황

동별 건물좌표 클러스터링



# 05 결론

01 공개 데이터

02 비공개 데이터

## 군집화 대상 좌표 리스트 업데이트

세대수 \* 5

분리수거장 \* 3

장려금 \* 2



- 쓰레기 발생량에 가장 큰 영향을 주는 세대수(인구수)에  
가장 큰 **가중치 5**
- 기존 분리수거장이 있는 공동주택 위치에 **가중치 3**
- 장려금을 받고 있는 공동주택에 **가중치 2**

곱해줌으로써 좌표 리스트 업데이트

# 05 결론

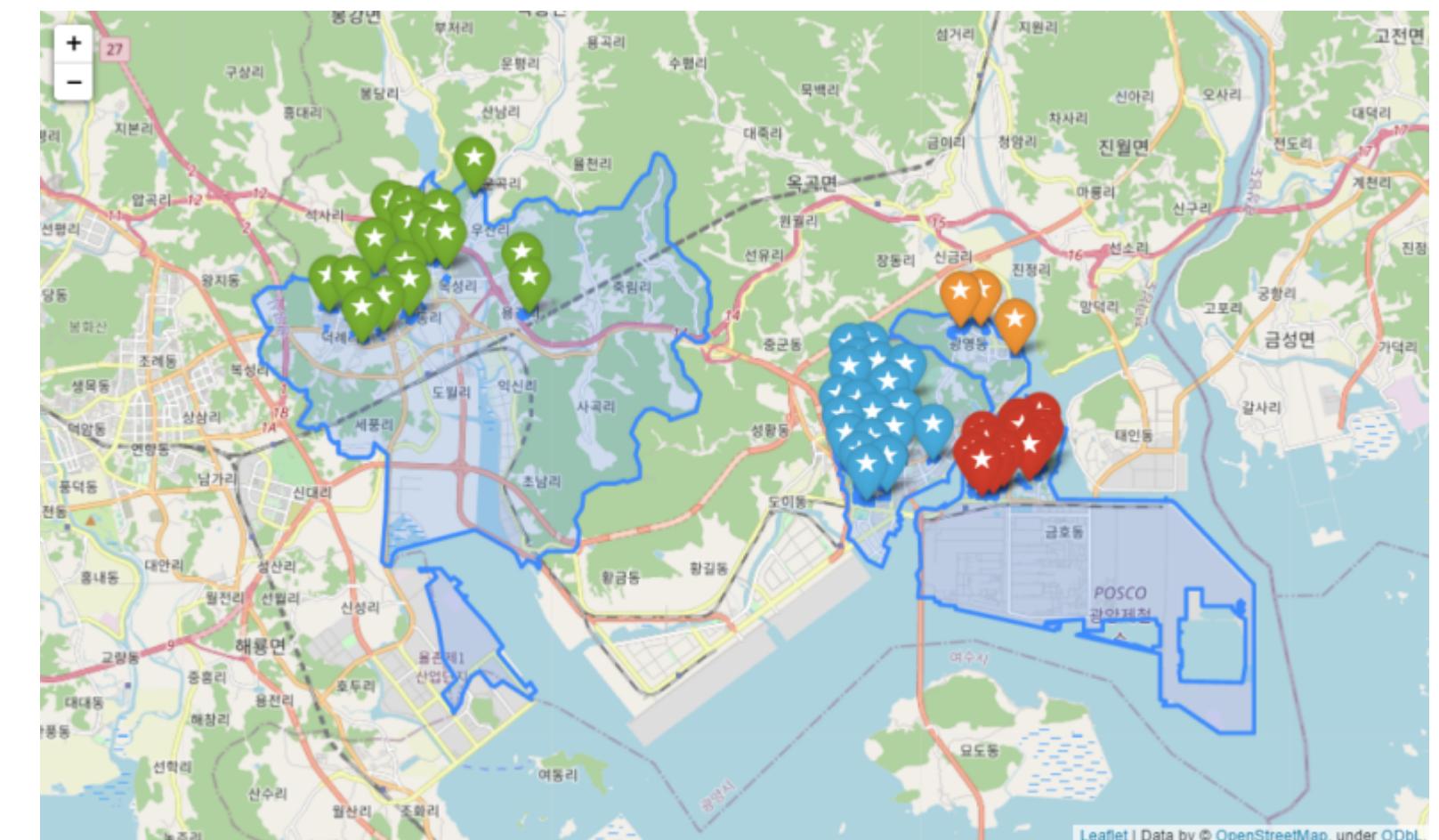
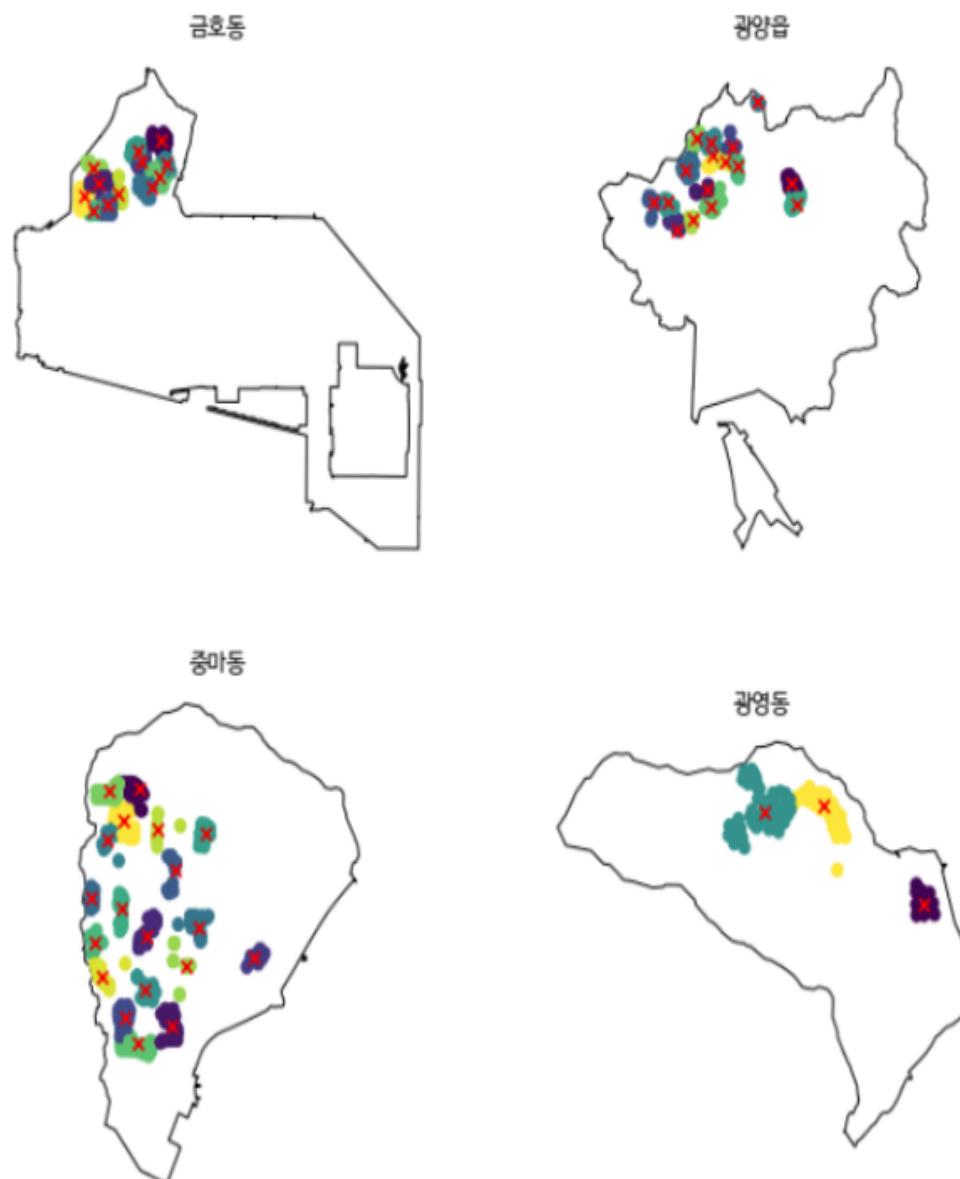
01 공개 데이터

02 비공개 데이터

## 개선된 모델 사용

클러스터링 결과 행정구역별 수거기 최종 위치 선정

동별 건물좌표 클러스터링



광양읍 16개

금호동 12개

중마동 19개

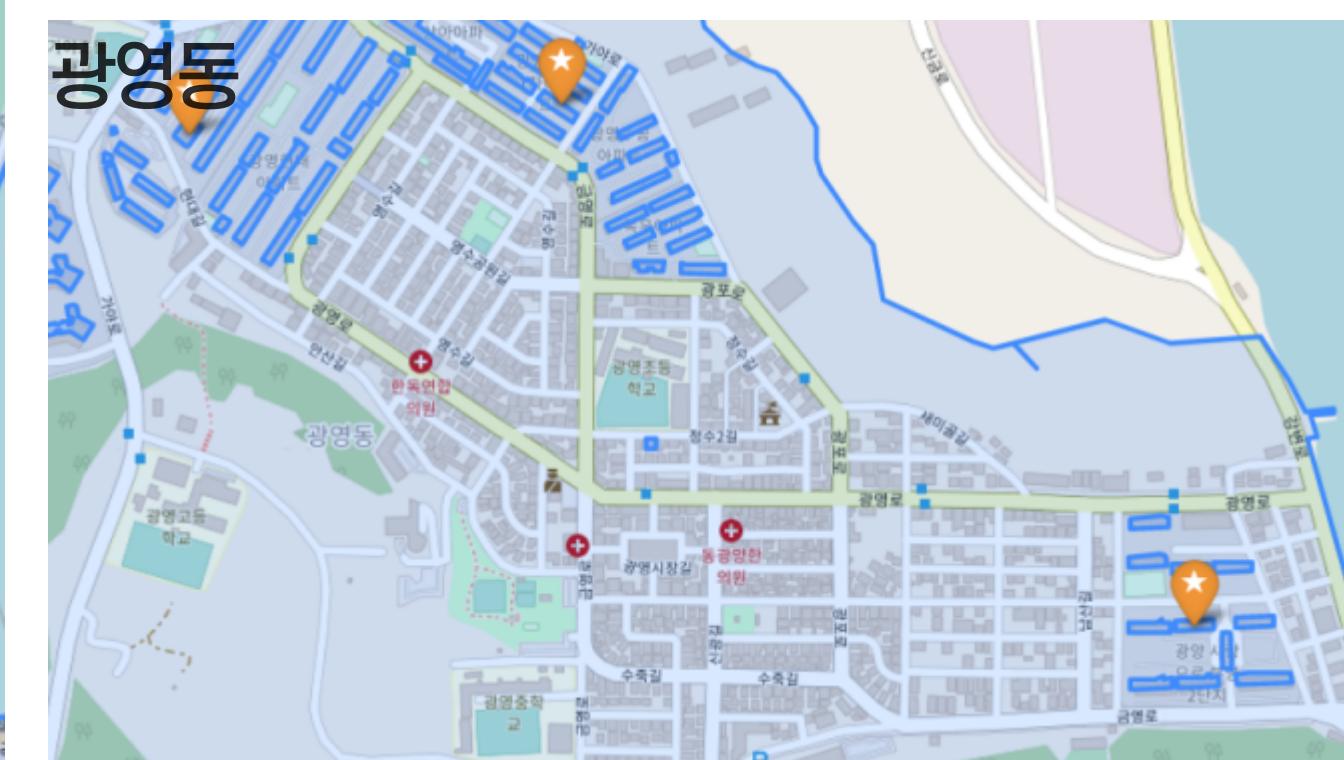
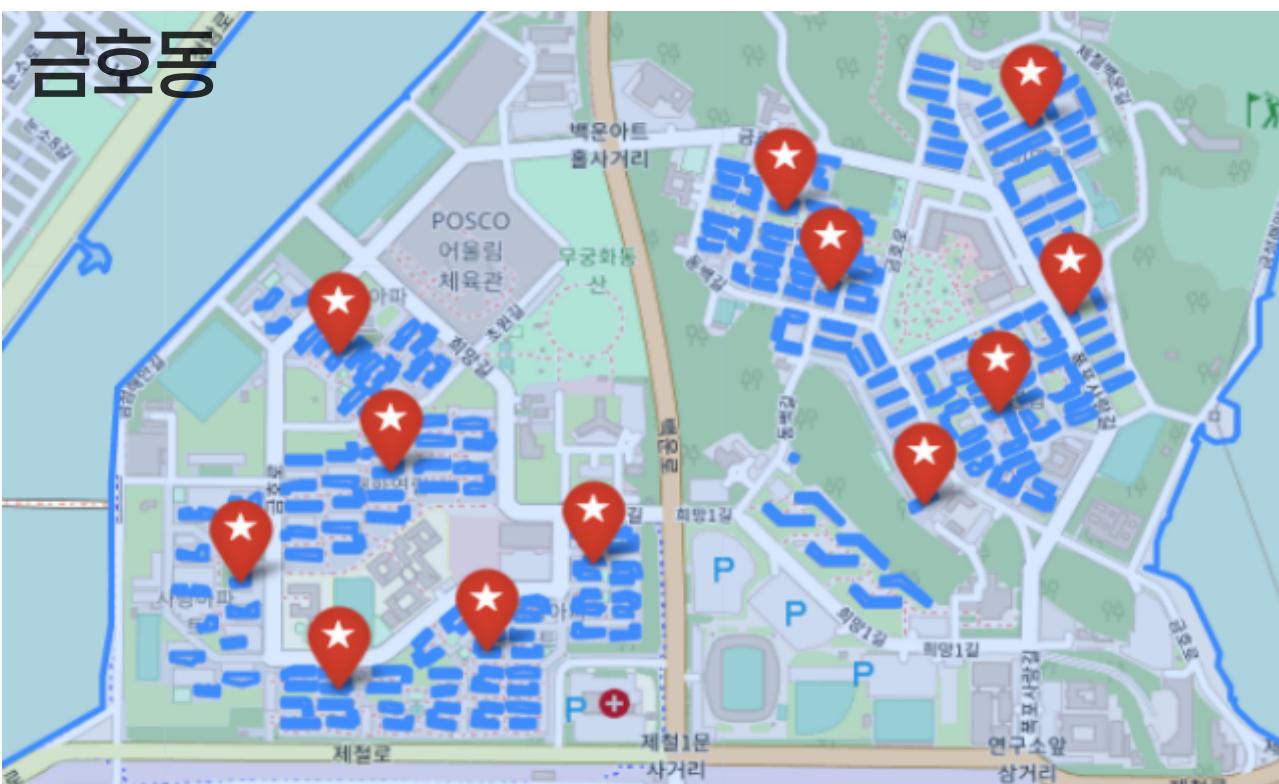
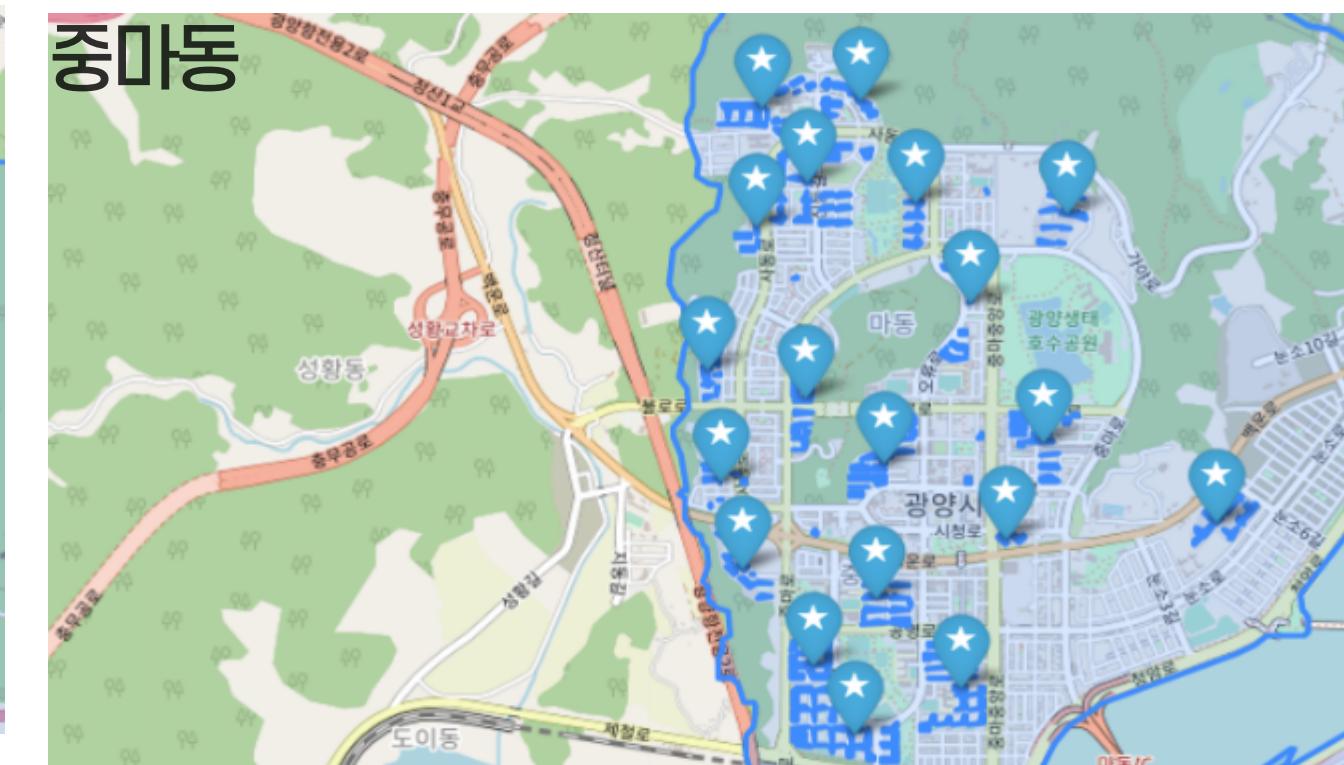
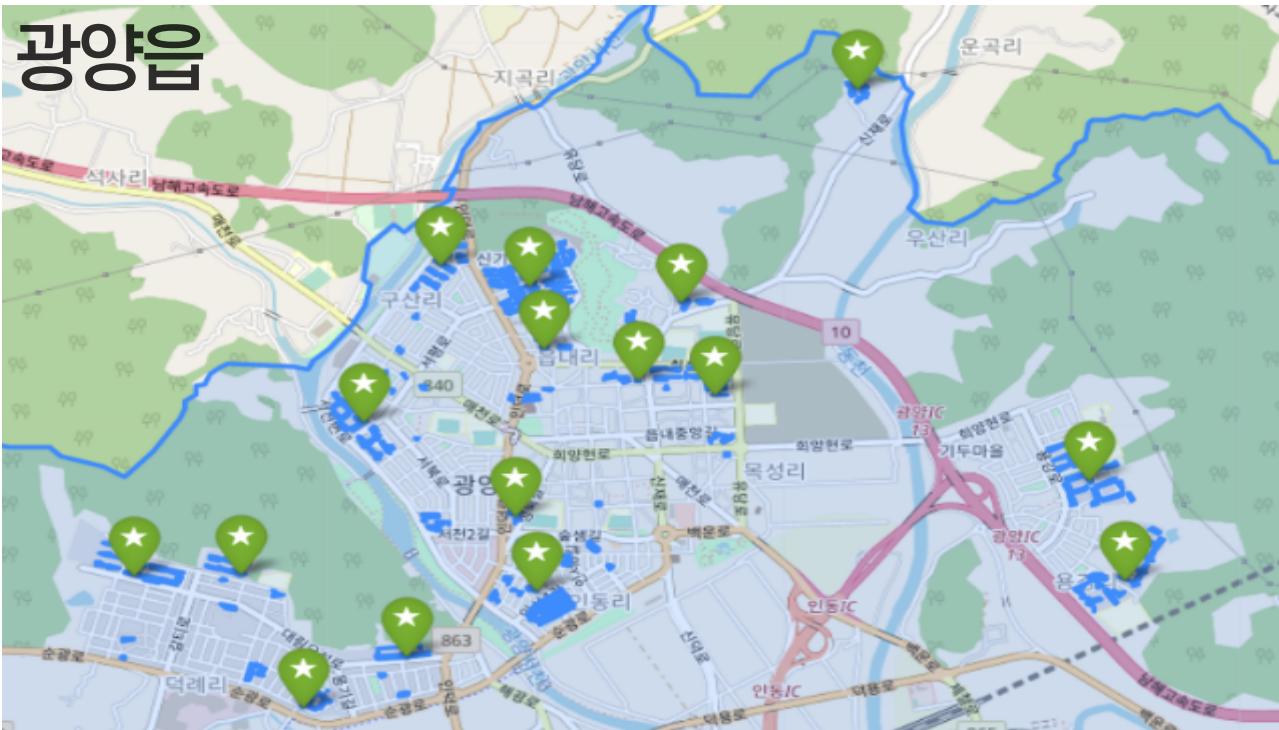
광영동 3개

# 05

# 01 공개 데이터

## 02 비공개 데이터

# 비공개 데이터 표



06

요약 및 한계

01. 요약
02. 한계

# 06

## 요약 및 한계

### 01 요약

### 02 한계

# 요약

## 결과 유의성

## 데이터 활용성

## 분석 창의성

※ 쓰레기 발생에 영향을 미치는 특성 중요도를 파악하여 최적의 위치에 잘 배치됨  
- 행정구역별 특성을 파악하고 그에 따른 개수를 선정, 가중치를 통한 최적의 위치 할당

※ 제공된 데이터 중 자동수거기 최적 위치 선정과 관련된 풍부한 자료를 이용  
- 인구수, 세대수, 인구밀도, 장려금, 분리수거장 위치 등을 활용  
- 공동주택에 해당되는 구역 중 인구 분포가 낮은 구역을 필터링하여 최적의 설치 개수 선정

※ 다양한 회귀 모델을 사용하여 행정구역에 따른 쓰레기 발생량 증가를 예측  
※ 교차검증, 실루엣 스코어 등 여러 평가지표를 활용해 검증 신뢰도를 향상시킴  
※ 장려금, 분리수거장, 세대수 데이터를 활용해 적절한 임의의 가중치 부여

# 06 요약 및 한계

## 01 요약

## 02 한계

# 요약

## 경제성

※ 추가 자동수거기 설치 계획서 관련 자료 참고하여 빠르게 자동수거기 위치 도출 가능  
- 설치할 구역의 인구 분포 확인 필요

## 발전 방향

※ 공동 주택 내 주차장, 소방도로, 보행도로 등의 위치 데이터 확보 시  
더 정확한 최적 위치 도출 가능  
※ 공동 주택 내 주민들을 위한 편의 및 기타 시설 설립에 관련 자료 활용이 기대됨

# 06

## 요약 및 한계

01 요약

02 한계

# 한계

## 데이터의 특성으로 인한 학습의 본질적인 한계가 존재

- 1 데이터와 학습에 많은 비중을 차지했던 점은 인구수가 대부분이었기 때문에, 그렇지 않은 지역들에 대한 학습이 부족함
- 2 제공된 데이터와 정보가 세부적인 위치 선정에 대해 적합하지 않은 부분이 다소 있기에 어려움이 있음

따라서 인구의 특성에 대해 세분화하고 지역 별로 적용시키며

실질적으로 수요가 높은 예상지역에

다른 특성의 가중치를 더하고 필터링하는 작업에 신경을 썼지만,

**특이성 및 창의성을 갖는 관점을 발견하지 못한 것에 아쉬움**

---

# 감사합니다

---

