

# 데이터 기반 선형공원 입지 선정 도시녹지 확충을 위한 최적 후보지 도출

#선형공원 #도시녹지확충 #데이터분석 #열섬해소

박유나

# 01. 서론

## 선형공원이란?

선형공원은 길고 좁은 띠 형태로 조성된 공원으로, 기존의 도로, 철도, 수로, 고가도로, 폐선부지 등을 활용해 만든 연결형 녹지공간입니다. 보행자와 자전거 중심의 이동 및 여가활동이 가능한 도시 속 그린 인프라로 기능합니다.

경의선숲길



출처: 도시정비뉴스 「연남동 경의선숲길, 카페·음식점 상권 확장된다」

양재천 생태공원



출처: 연합뉴스 「서초구 양재천 복원사업, 英 '그린월드어워즈' 은상 수상」

하이라인 파크



출처: NewYork.kr 「뉴욕 하이라인 파크」

크레센트 그린웨이



출처: realestate.com.au 「Cynthea Teague Crescent, Greenway」

## 선형공원의 잠재력?

선형 자원은 철도, 고가도로, 수로 등 도시 곳곳에 존재하는 길고 단절된 유휴 인프라입니다.

이러한 공간은 선형공원으로 전환될 경우, 열섬 완화, 보행 회복, 지역 활성화 등 도시 문제를 통합적으로 해결할 수 있는 높은 공간 잠재력을 지닙니다.

## 02. 문제 배경 및 필요성 분석

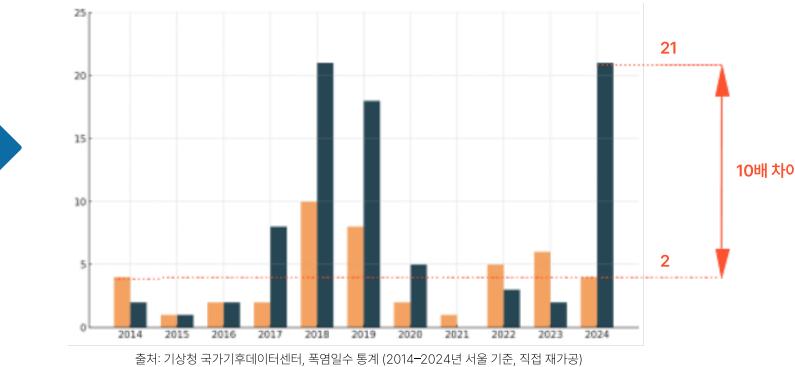
### 녹지의 역할

식물의 증산작용이 공기를 식히고  
토양·그늘이 지면 온도 상승을 억제함

### 녹지 불균형 = 열 축적 불균형

도심부는 녹지가 부족하고  
아스팔트, 콘크리트 등 불투수면 비율이 높음  
-> 낮 동안 열을 흡수하고 밤에도 열을 방출하지 못함

[ 10년 서울 7 ~ 8월 폭염 일수 추이 ]



서울시의 8월 폭염일수가 지난 10년 사이 10배가량 증가

뉴스1

### 서울 35도 넘긴 날 9→58일로 늘어... '열섬효과'로 대도시 폭염 위험

"특히 지난 2018년엔 35도를 넘는 날이 21일을 기록했는데  
이는 앞서 10년간의 일수를 합친 것보다 많은 수준이다."

서울시의 열섬 악화

▶ 하지만 서울 도심은 빽빽한 건물과 도로, 철도로 녹지 확보가 어려움

▶ 제한된 공간을 활용하기 위해서는 선형 자원을 활용한 '선형 공원'을 조성

### 선형 자원이란?

선형 자원은 철도, 고가도로, 수로 등 도시 곳곳에 존재하는 길고 좁은 유휴 인프라입니다.

### 선형 공원이란?

이러한 선형 자원을 따라 녹지를 연결·확장하여 만든 공원으로,  
열섬 완화·보행 환경 개선·지역 활성화 등 복합적 도시 문제를 해결할 수 있는 잠재력을 지닙니다.

경의선숲길



출처: 도시정비뉴스 '연남동 경의선숲길, 카페·음식점 상권 확장된다'

양재천 생태공원



출처: 연합뉴스 '서초구 양재천 복원사업, 英 '그린월드어워즈' 은상 수상'

하이라인 파크



출처: NewYork.kr '뉴욕 하이라인 파크'

### 분석 목표

행정동별 도시환경·녹지여건·활동특성을 분석하여, 선형공원 조성이 시급한 우선 입지 행정동 선정

# 03. 분석 프레임워크

## 분석 로드맵

데이터 수집 및 전처리

### 1 녹지 부족도

- 세대당 공원면적
- 공원 비율

### 2 접근성

- 지하철 밀도
- 버스 밀도
- 인구 1만명당 지하철역 수
- 인구 1만명당 버스정류장 수

### 3 사회 수요

- 인구 밀도
- 유동 인구
- 집객시설 밀도

### 4 경제·부지 여건

- 5년 평균 공시지가
- 매출 밀도

행정동 단위로 통합·정제하여 분석 기반 구축

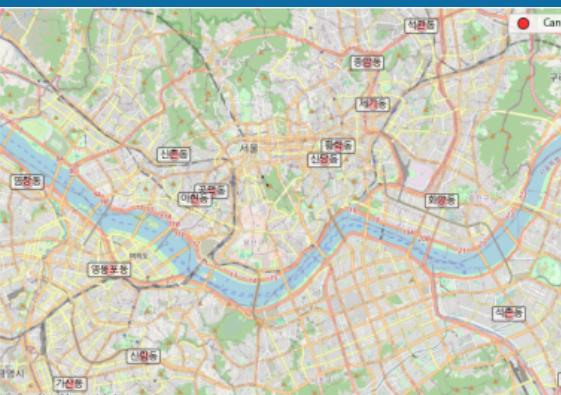
### 클러스터링 - 후보 행정동 선정

#### 최종 변수 선정

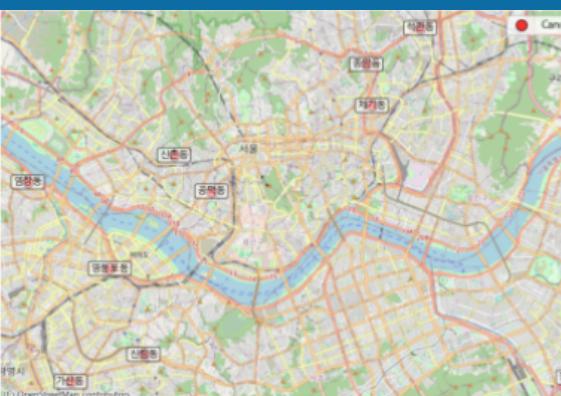


$Z_i$  (행정동의 실제 데이터 값)

#### Agglomerative



#### K-Means



### 휴리스틱 알고리즘 - 최적 행정동 선정

#### 변수 중요도 설정 (AHP 적용)

각 변수의 중요도는  $CV \times IG$ 로 계산  
→ 중요도 비율로 쌍대비교 행렬 구성  
→ 고유벡터로 가중치 도출 (AHP 방식 유사 적용)  
→ 고유벡터로부터 가중치  $w_1 \sim w_5$  도출

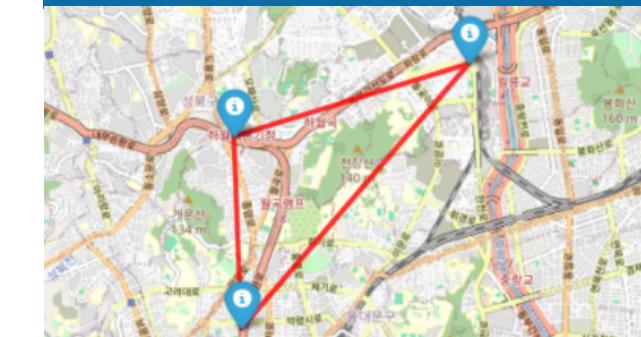
$$Q(x) = \sum (w_i \cdot Z_i), i = 1 \text{ to } 5$$

#### INTER 지수 계산

$$INTER(i) = \frac{IC_i \cdot OA_i}{IA_i \cdot OC_i}$$

중심지와의 유입·유출 통행량을 기준으로  
해당 지역의 중심성과 연결성을 수치화  
값이 클수록 → 중심성과 상호작용이 강한 지역

#### 거리 및 외곽 조건 기반 휴리스틱 최적화

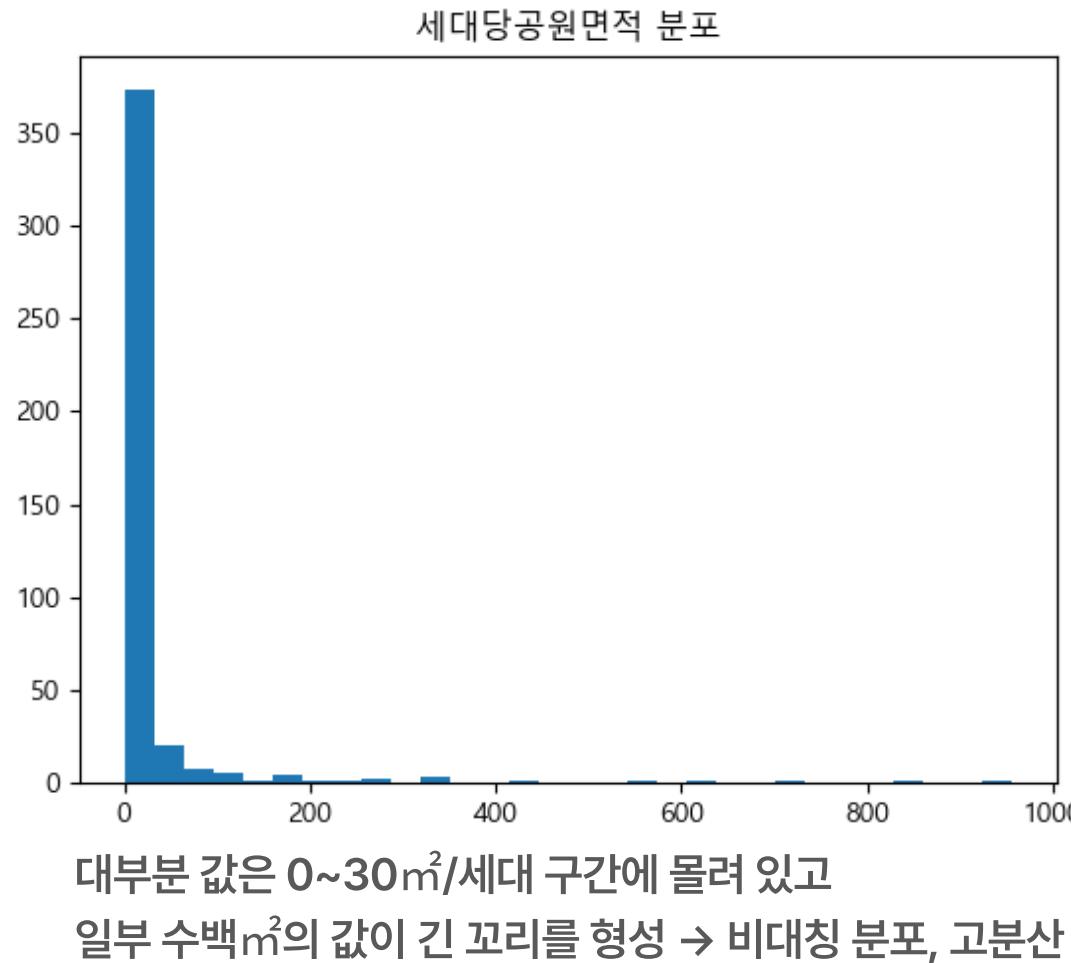


결과 → '제기동', '종암동', '석관동'

## 03. 분석 프레임워크

### 녹지 부족도에 대한 변수

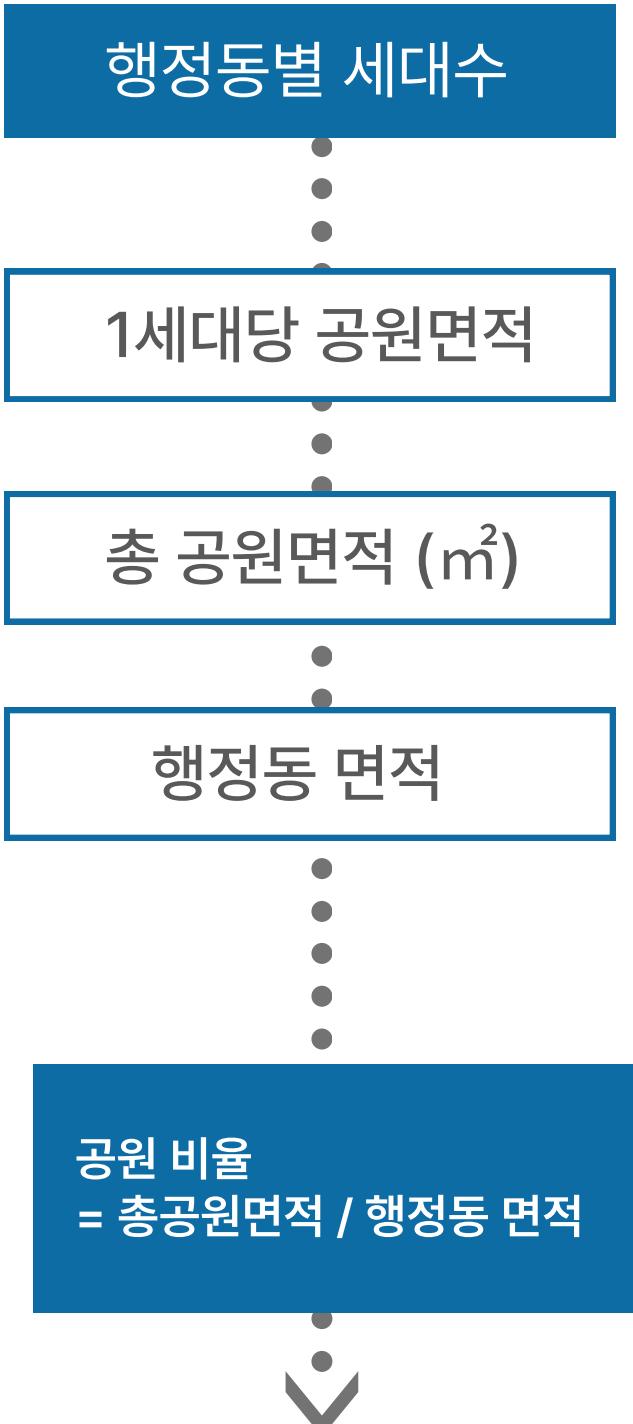
#### - 세대당 공원면적



모델용 피처로는 로그값 활용 / 시각화·정책 해석에는 원본값 유지  
→ 이중 전략으로 해석력과 모델 안정성 동시 확보

처리방식	
1세대당 공원면적	np.log1p 적용 (로그 변환)
공원값의 극단치 대응	로그 변환으로 모델 안정성 확보
정보값 왜도/왜곡 해결	히스토그램 기준 분포 왜곡 완화
로그값 활용 목적	모델 성능 및 해석력 개선
원본값 활용 목적	보고서·시각화에는 원본값 사용이 더 직관적

#### - 공원 비율

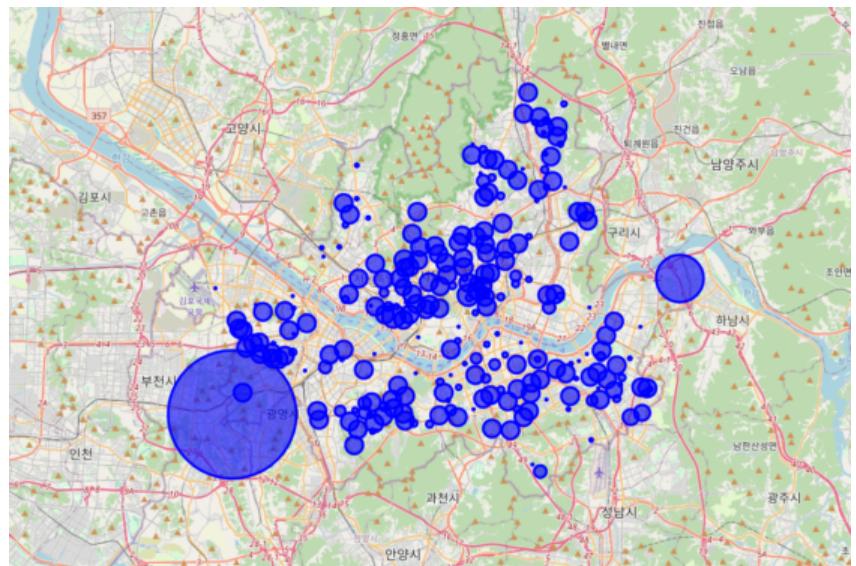


## 03. 분석 프레임워크

### 접근성에 대한 변수

#### - 지하철 접근성

- 지하철역 위경도 확보
- 행정동 경계와 공간 결합
- 행정동별 지하철역 수 계산

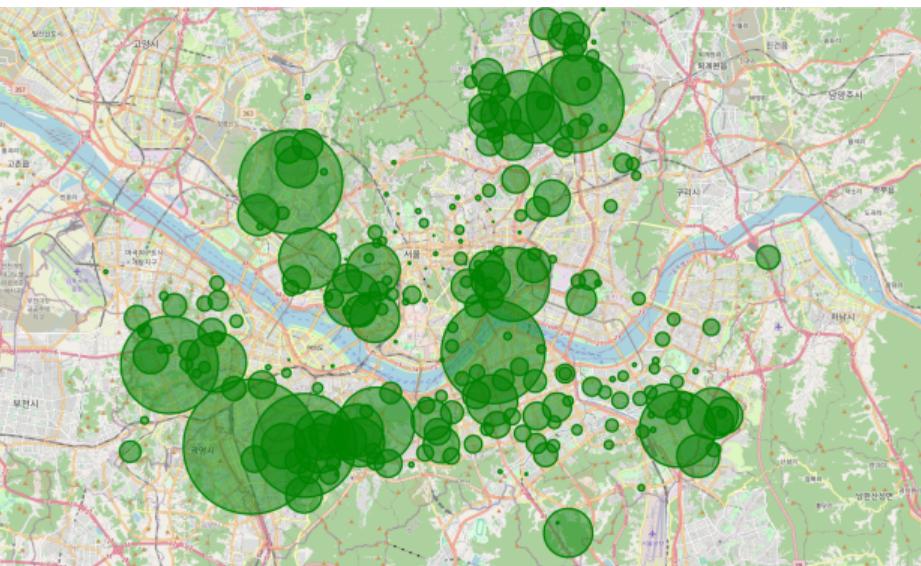


행정동별 지하철역 밀도 시각화 ( $\text{역}/\text{km}^2$  단위)

- 밀도 기반 접근성
- 인구 보정 접근성

#### - 버스 접근성

- 버스 정류장 위경도 확보
- 행정동 경계와 공간 결합
- 행정동별 버스정류장 수 계산

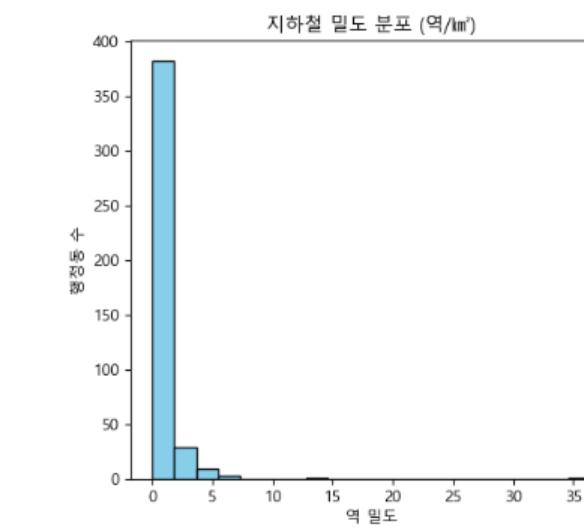


행정동별 버스정류장 밀도 시각화 ( $\text{정류장}/\text{km}^2$  단위)

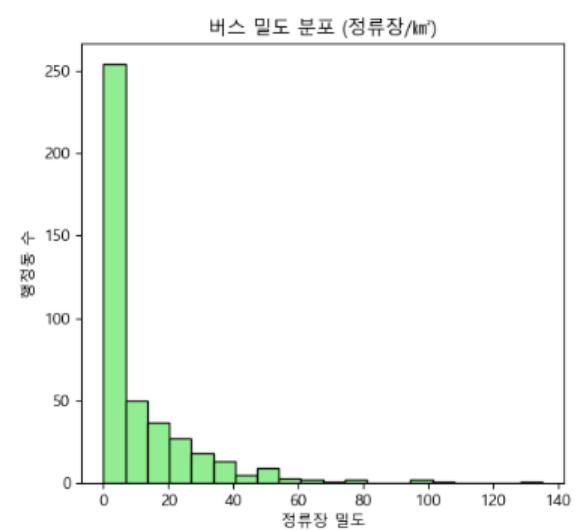
- 밀도 기반 접근성
- 인구 보정 접근성

#### - 인구당 접근성

- 행정동별 총 인구 확보
- 지하철역 수  $\div$  인구  $\times 10,000$
- 버스정류장 수  $\div$  인구  $\times 10,000$



지하철 밀도 분포 (히스토그램) / 버스 밀도 분포 (히스토그램)



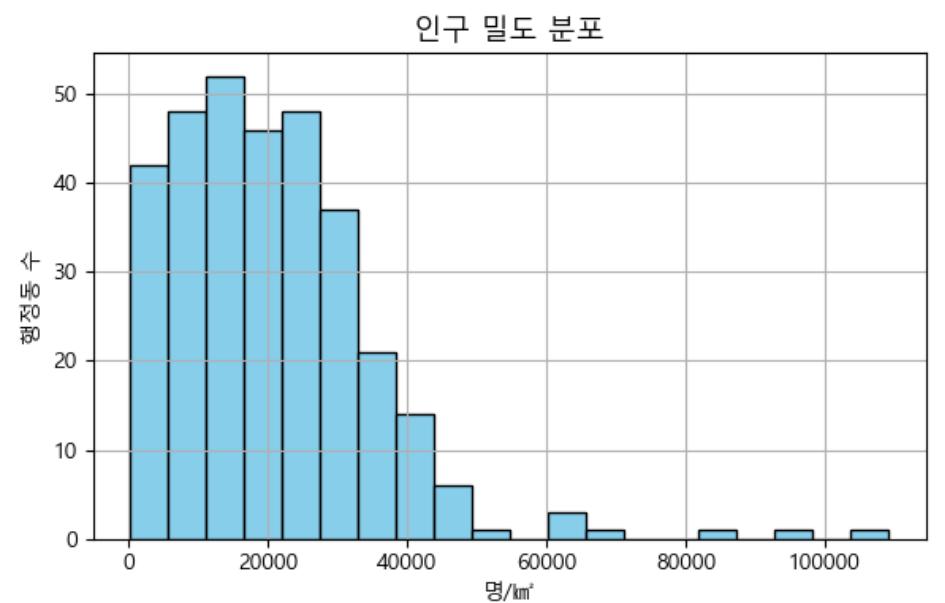
→ 두 지표 병행 사용으로 공간적·생활적 접근성 모두 고려

## 03. 분석 프레임워크

### 사회 수요에 대한 변수

#### -인구 밀도

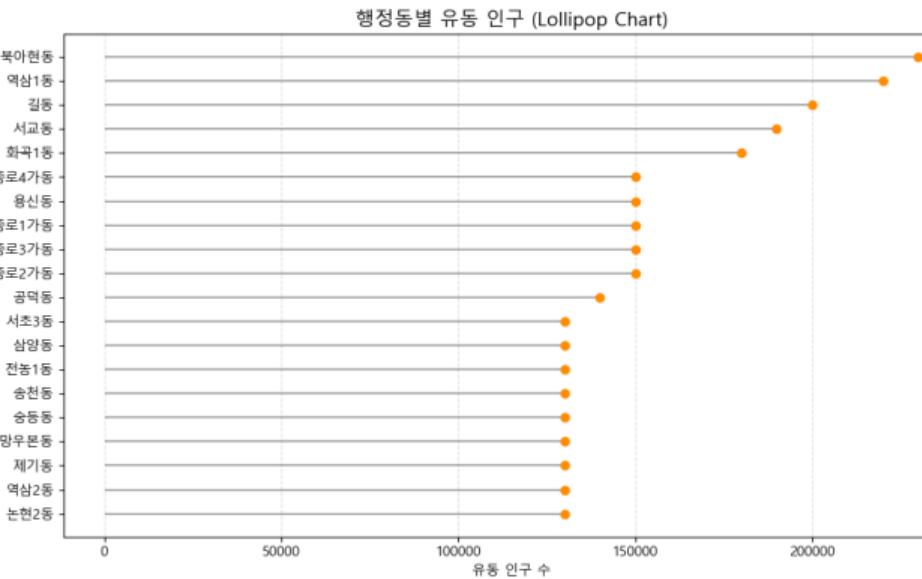
- 행정동 단위 인구 데이터 활용
- 행정동 기준 공간 결합
- 행정동 인구 / 행정동 면적



행정동 면적당 인구 수 분포 (명/km<sup>2</sup> 기준, 정규성 확인 목적)

#### - 유동 인구

- 시간대별 유동인구 데이터
- 평균값(시간·요일 기준) 계산



상위 행정동 유동 인구 비교 (시간·요일 평균 기준 체류 수요)

#### - 집객시설 밀도

- 싱글 단위 집객시설 데이터
- 행정동 기준 공간 결합
- 행정동 집객시설 수 / 행정동 면적



행정동별 집객시설 밀도 비중 시각화 (면적당 주요 시설 수 Top 25)



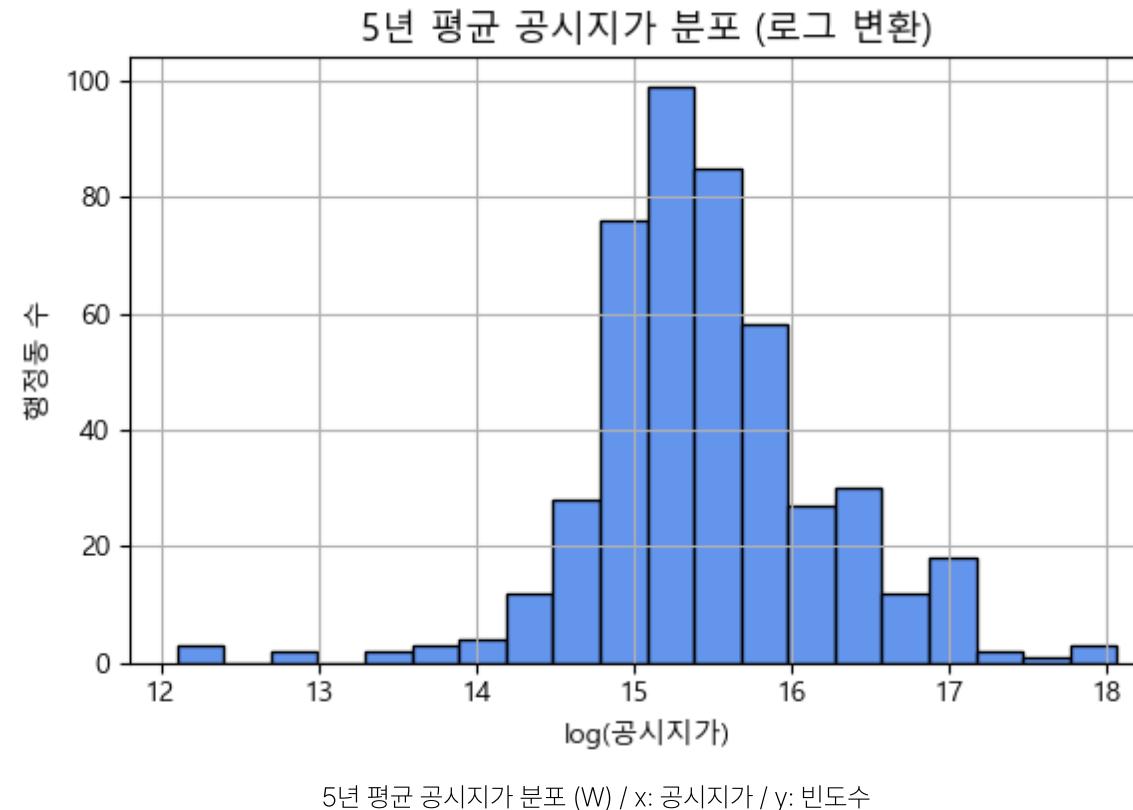
모든 변수 행정동 면적 기준으로 밀도화 → 정적 + 동적 + 기능적 사회 수요 반영

## 03. 분석 프레임워크

### 경제부지 여건에 대한 변수

#### - 5년 평균 공시지가

- 공시지가(W) 로그 변환 (정규성 확보)
- 행정동별 면적당 평균값 활용



#### - 추정 매출액

- 공시지가(W) 로그 변환 (정규성 확보)
- 행정동별 면적당 평균값 활용

추정매출액/변환값	
사직동	561,650
상정동	92,038
청운효자동	433,498
평정동	135,760
부양동	142,144
가회동	184,030

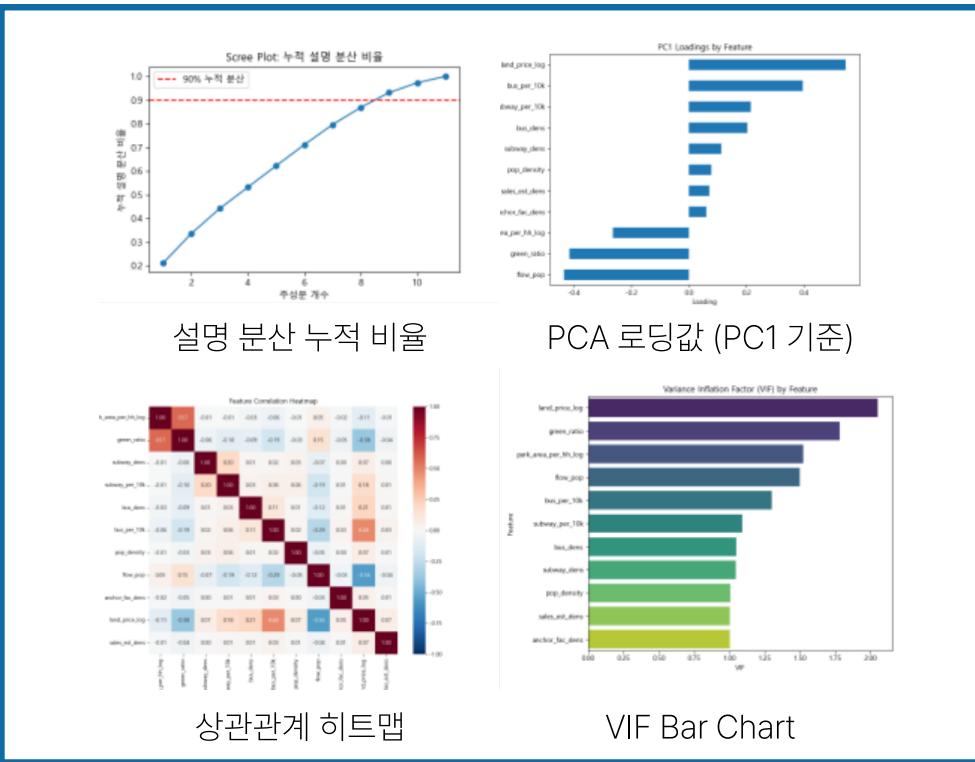


모델용 피처로는 로그값을 추가하고 / 보고·시각화용 정제 해석에는 원본값 유지

# 04. 1차 분석: 군집 기반 후보지 도출

## 1. 클러스터링에 쓰일 변수 선정 및 전처리

No.	활용 가능한 변수
1	세대당 공원 면적
2	공원 비율
3	지하철 밀도
4	버스 밀도
5	인구 1만명당 지하철역 수
6	인구 1만명당 버스정류장 수
7	인구 밀도
8	유동 인구
9	집객시설 밀도
10	5년 평균 공시지가
11	매출 밀도



### 상관관계 히트맵

→  $|상관계수| \geq 0.8$ 인 변수 쌍 중 하나 제거 (중복 정보 제거 목적)

### VIF (Variance Inflation Factor)

→  $VIF \geq 10$ 인 변수 제거 (다중공선성 우려)

### PCA 로딩값 (PC1 기준)

→ PC1 기준 기여도가 낮은 변수 제거 (절댓값 하위 50%)

### 분산 임계값 (Variance Threshold)

→ 분산  $< 0.5$ 인 변수 제거 (정보량 부족)

## [ 최종 변수 선정 ]



### Z-score 표준화

$$Z = \frac{X - \mu}{\sigma}$$

X: 원래의 데이터 값

$\mu$ : 해당 변수의 평균

$\sigma$ : 해당 변수의 표준편차

Z: 표준화된 값 (Z-score)



서로 다른 단위와 범위를 가진  
변수들을 동일한 기준으로 맞추기 위해

# 04. 1차 분석: 군집 기반 후보지 도출

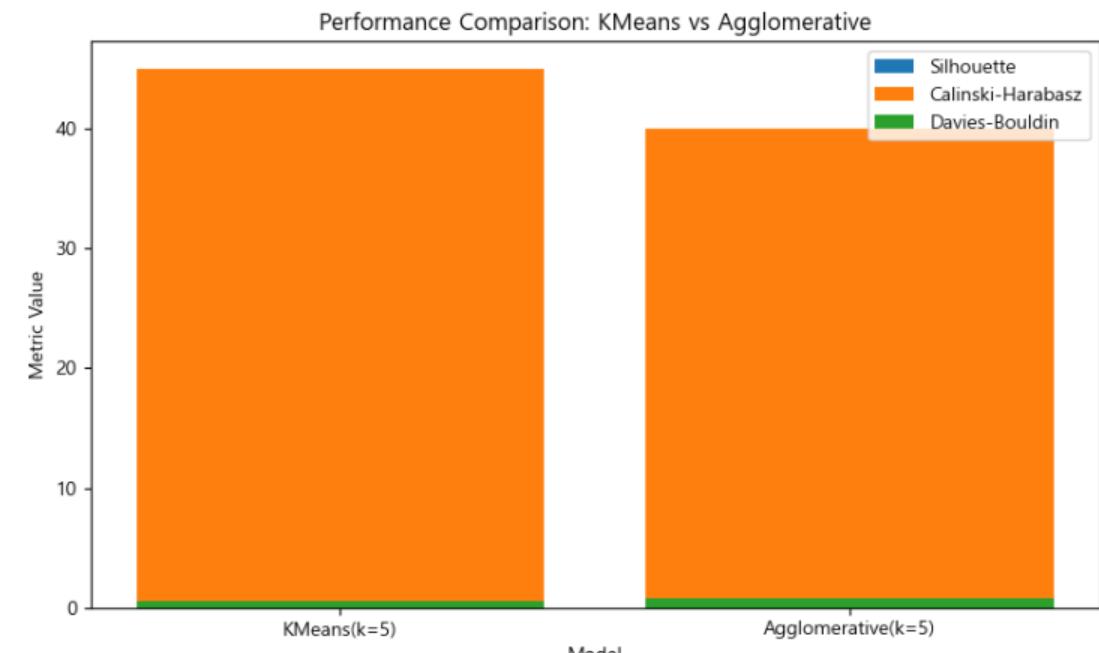
## 2. 클러스터링을 통한 후보 행정동 선정

### K-means

군집의 중심점을 기준으로 데이터를 나누는 비계층적 중심 기반 클러스터링 기법

### Agglomerative

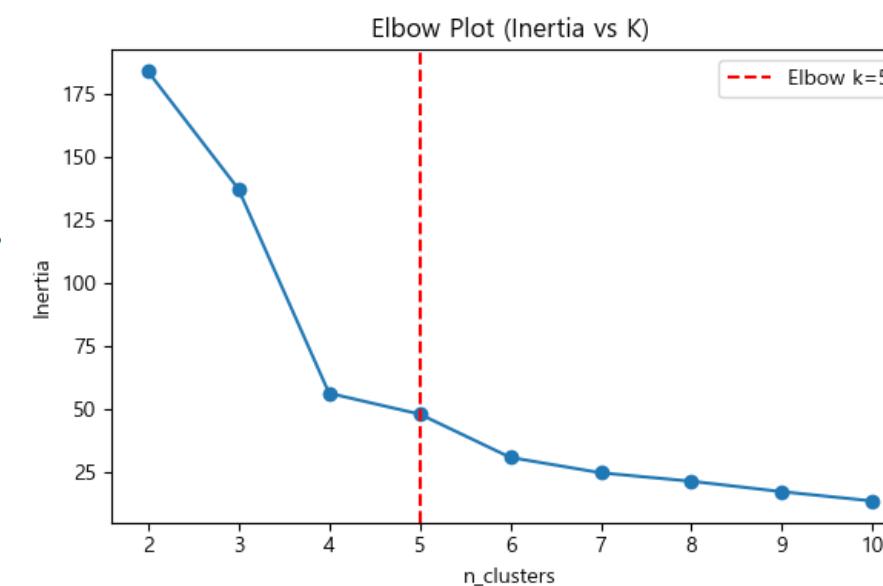
데이터 간 거리 정보를 기반으로 점차 병합해 나가는 계층적 병합형 클러스터링 기법



Silhouette, Calinski-Harabasz, Davies-Bouldin 지표를 시각화하여 비교한 결과,

K-means가 전반적으로 더 우수한 군집 품질을 보여

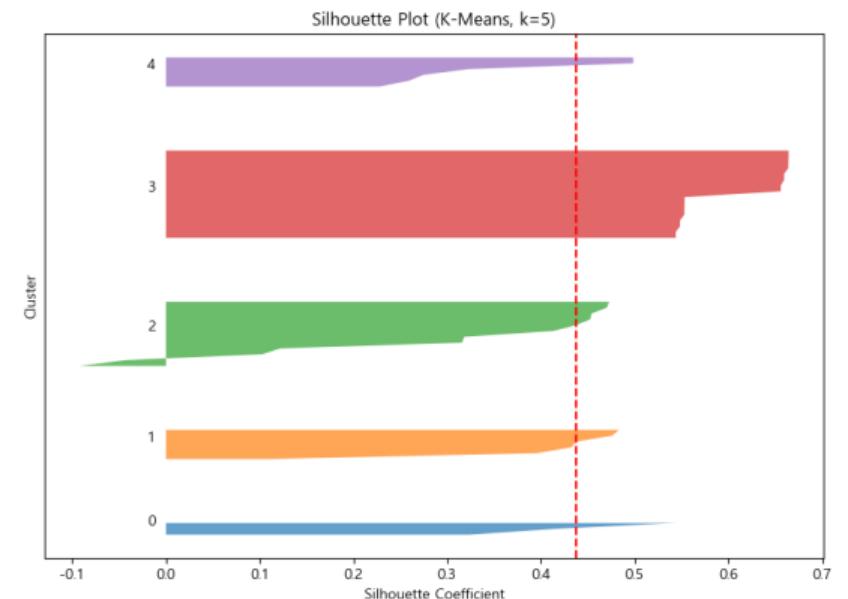
**K-means 클러스터링**을 최종 선택



### 엘로우 그래프

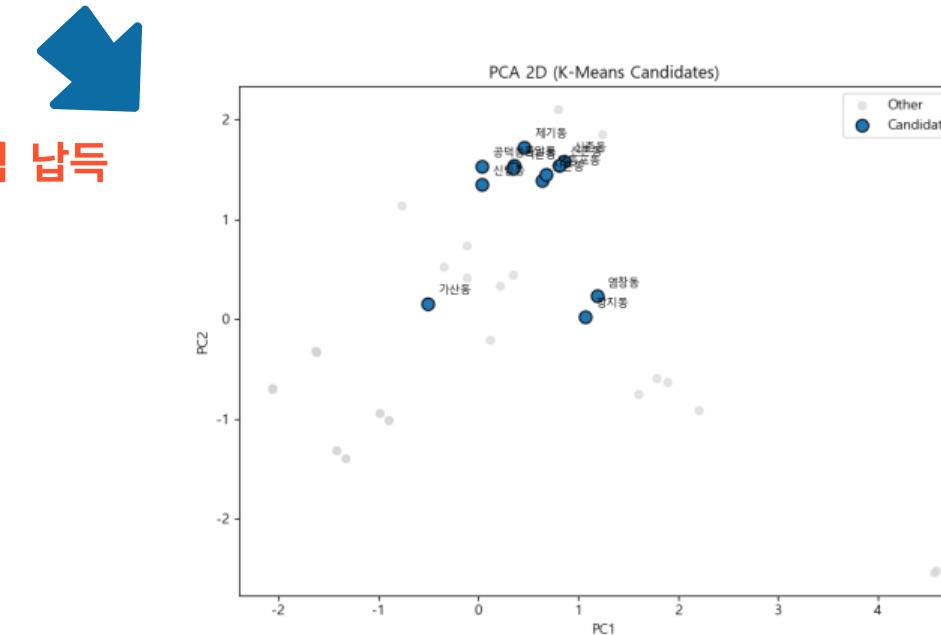
K=5에서 기울기 감소 → 최적 k 후보 도출

수치적 검증



### 실루엣 점수 분포

Elbow로 제안된 k=5에 대해, 실루엣 점수를 통해 군집의 응집도와 분리도를 수치적으로 검증



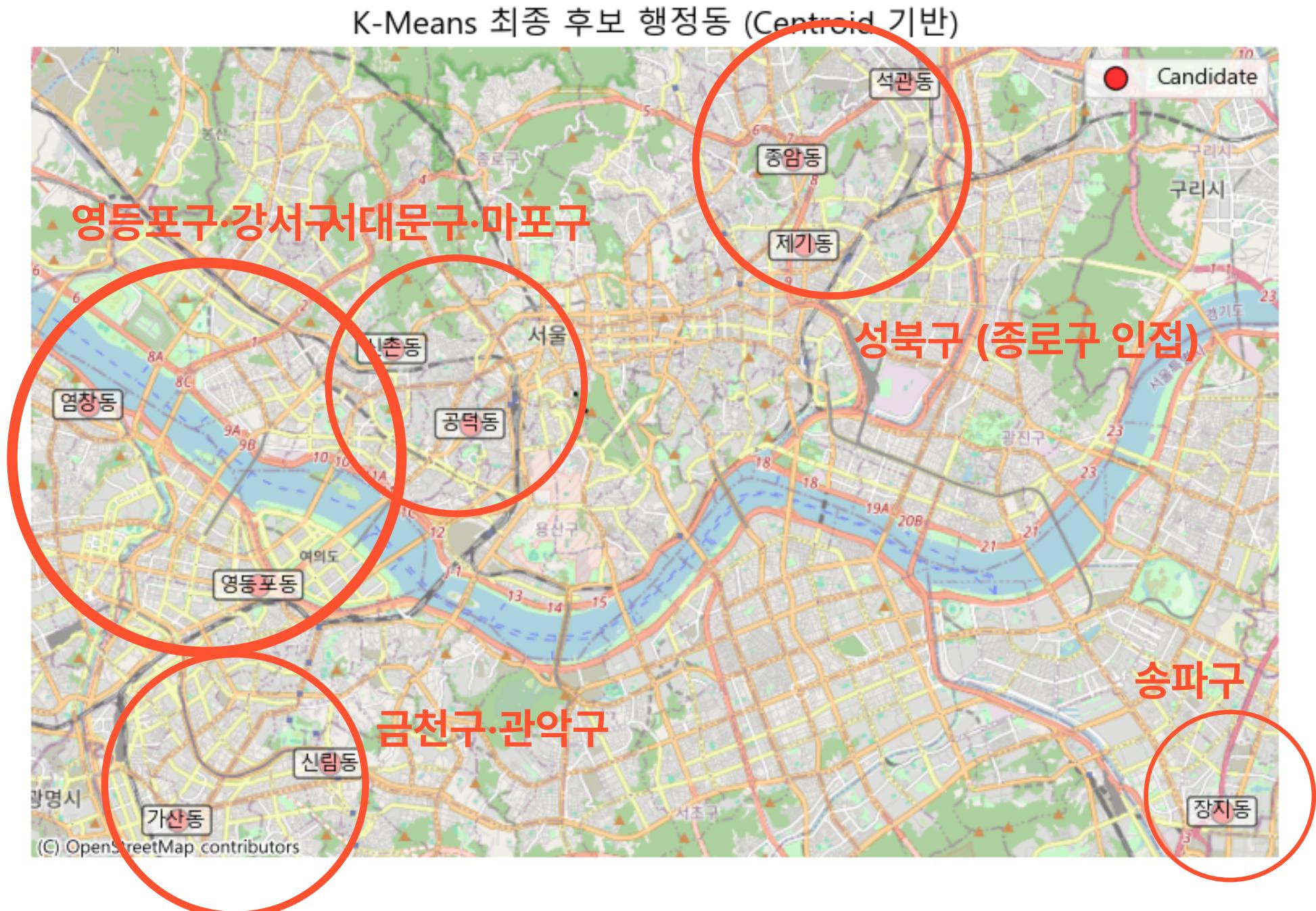
### PCA 시각화

K=5로 군집화한 결과를 2D로 투영

→ 후보지들이 유사한 특성을 기반으로 응집된 군집을 형성함을 시각적으로 확인

## 04. 1차 분석: 군집 기반 후보지 도출

### 3. K-means 클러스터링 결과 시각화 및 후보 행정동 선정



**최종 선정된 후보 행정동 (K-means 기반)**

**최종 선정된 후보 행정동 (K-means 기반)**

K-means 클러스터링 결과, 아래 15개 행정동이  
공원면적, 유동인구, 교통 접근성, 공시지가 등에서 유사한 특성을 공유하며  
선형공원 조성의 우선 검토 대상지로 선정되었습니다.

- 성북·종로권역: 제기동, 종암동, 석관동
- 서북권: 신촌동, 공덕동, 서대문구·마포구 일대
- 영등포권: 영등포동, 양화동, 창동
- 동남권: 장지동, 송파구
- 도심외곽·연계지역: 신림동, 가산동, 금천구·관악구 일대

이들 지역은 동일 클러스터 내에서도 지리적 근접성과 기능적 연계성이 높아  
"선형공원 도입 시 시너지 효과"가 클 것으로 기대됩니다.

# 05. 2차 분석: 휴리스틱 기반 입지 평가

## 4. 후보 행정동 중 최적 입지 선정을 위한 휴리스틱 기반 점수 산정

### (1) 변수 중요도 산정 (AHP 기반 Q(x) 점수 도출)

목적

각 행정동의 특성을 종합적으로 고려해 중요도 기반 점수를 부여  
→ 녹지 부족, 접근성, 사회 수요, 경제 여건 모두 반영

변동 계수(CV) 계산 각 변수의 상대적 변동성 측정

$$CV_i = \frac{\sigma_i}{\mu_i}$$

정보이득(IG) 계산 변수의 정보 제공력 (엔트로피 기반)

$$IG_i = 1 - \frac{H(X_i)}{H(X)}$$

중요도 점수 산출 두 지표를 곱하여 변수별 중요도 결정

$$\text{중요도}_i = CV_i \times IG_i$$

AHP 고유벡터 도출 중요도 기반 쌍대 비교 행렬 구성 → 고유벡터 w 도출

최종 Q(x) 점수 산출 행정동 단위 z-score 정규화된 값  $Z_i$ 에 대해:

$$Q(x) = \sum (w_i \cdot Z_i)$$

→ 모든 변수에 대해 가중합으로 행정동 종합 점수 Q(x) 부여

→ Q(x) 상위 후보지만 우선 선정

### (2) 최종 입지 조합 평가 (INTER 기반)

#### INTER 지표 정의

중심지역과 후보지역 간의 상호작용 비율 지표  
각 행정동의 중심지 유입·유출 교통량을 반영

$$\text{INTER}_i = \frac{IC_i + OC_i}{IA_i + OA_i}$$

- $IC_i$ : 중심지역 → i로 유입
- $OC_i$ : i → 중심지역으로 유출
- $IA_i$ : 전체 → i로 유입
- $OA_i$ : i → 전체로 유출

### (3) 최적화 수식 (조합 탐색)

목적

조합의 연결성과 Q(x) 조건을 함께 고려해 최적 후보지 조합 도출

$$\min Z = \sum (INTER_{ij} \cdot w_{ij}) - \alpha \cdot (\sum Q_{outer} - \sum Q_{adjacent})$$

- $w_{ij}$ : 후보지 간 연결 여부 ( $1/\text{INTER}$ )
- $Q_{outer}$ : 도심 외곽 후보의 Q(x) 총합
- $Q_{adjacent}$ : 인접 후보지들의 Q(x) 총합
- $\alpha$ : 조정 파라미터 (균형 조절용)

#### 제약조건

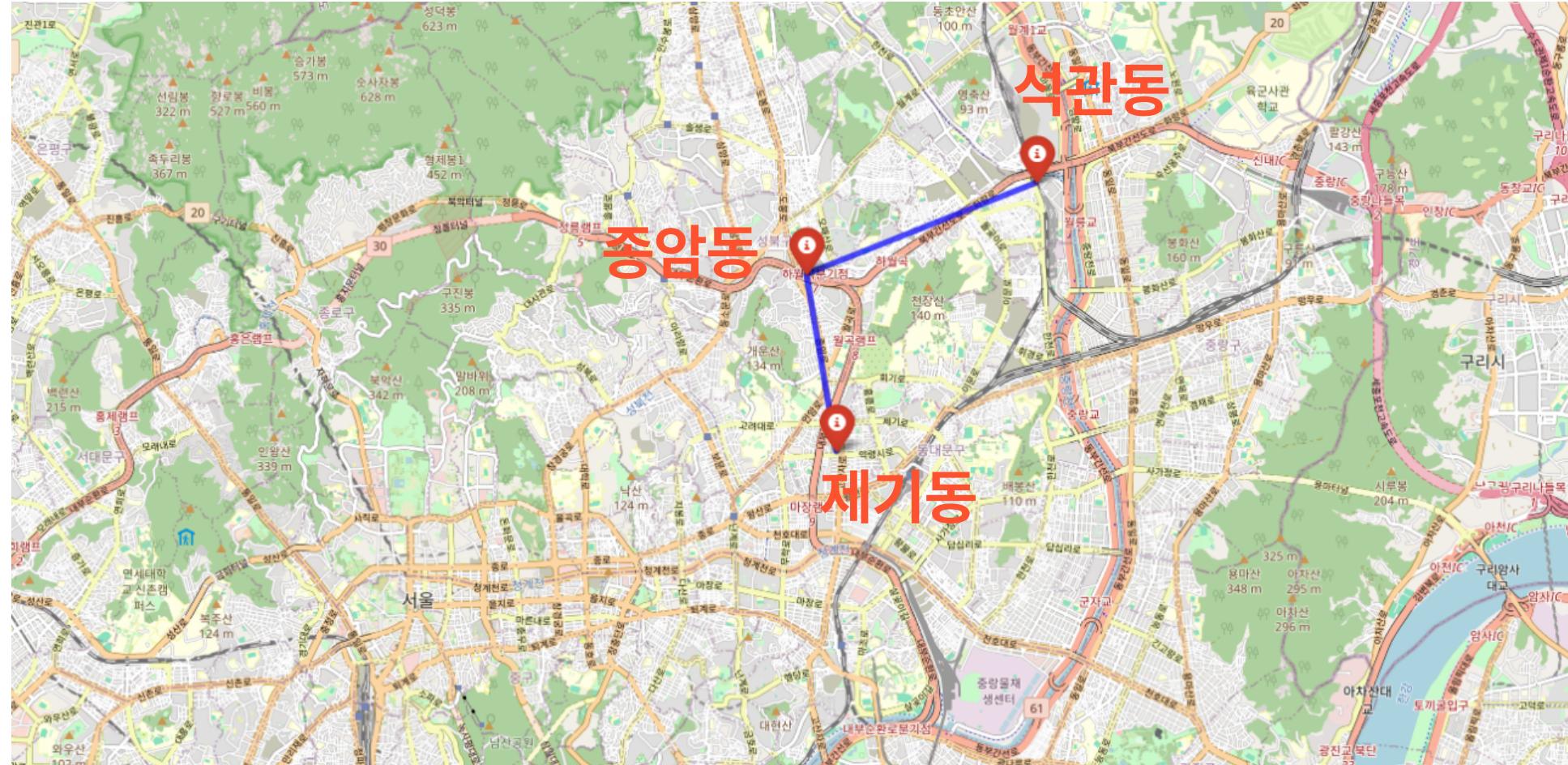
- 최소 1개는 외곽 후보 포함
- 총 연결수 제한
- 인접성 확보 (거리 조건  $\leq 5\text{km}$ )



결과 : Q(x) 기반 우선 후보지 중 → INTER 연결성 높은 조합을 최종 입지로 선정

## 06. 결론 및 인사이트 도출

### 5. 최적 행정동 선정 및 인사이트



### 프로젝트 전 과정 요약

- 도시환경, 녹지여건, 접근성, 사회 수요, 경제 여건 등 다양한 관점에서 총 11개 변수 수집 및 정제
- 변수 간 중복 제거 및 중요도 평가를 거쳐 K-means 기반 클러스터링으로 유사 특성 후보지 추출
- 이후 AHP 기반 점수화와 공간 연결성 분석(INTER)을 통해 데이터 기반 판단 + 공간적 타당성을 동시에 반영

→ 선형공원 조성이 시급하고 효과적인 최적 입지 3곳을 도출함

- 석관동: 인구 밀도↑·녹지율↓
- 종암동: 유동인구↑·교통접근성↑
- 제기동: 하천 인접·토지부담↓

데이터 분석을 통해 서울시에서 선형공원 조성이 시급한 3개 행정동(석관동, 종암동, 제기동)을 최종 추천합니다.

# 07. 부록

## 활용 데이터 목록 및 출처

활용 데이터셋명	출처
2020년 서울시 읍면동별 1세대당 도시공원 면적	산림빅데이터거래소
서울특별시_역사마스터 정보	서울 열린 데이터 광장
서울시 버스 정류소 위치 정보	서울 열린 데이터 광장
행정안전부_지역별(행정동) 성별 연령별 주민등록 인구수	서울 열린 데이터 광장
서울 생활 이동	서울 열린 데이터 광장
서울시 상권분석 서비스(집객시설-상권)	서울 열린 데이터 광장
서울시 상권분석 서비스(영역-상권)	서울 열린 데이터 광장
지역분석 개폐업수(률)	서울 열린 데이터 광장
서울시 번지별 공시지가 데이터	서울시 빅데이터 캠퍼스
서울시 행정동단위 분기별 추정매출액	서울시 빅데이터 캠퍼스
서울시 행정구역 (동별) 통계	서울 열린 데이터 광장
기상청_폭염일수	기상청 (국가기후데이터센터)

## 참고문헌

참고문헌명	저자 및 출처
AWS데이터를 활용한 서울시 도시열섬 분포의 변화 특성 분석 - 2005년도 및 2019년도 온도측정 자료 비교를 중심으로 -	박세홍 (중앙대학교) 배웅규 (중앙대학교)
The Social Equity of Urban Parks in High-Density Urban Areas: A Case Study in the Core Area of Beijing	배웅규 /Bae, Woongkyoo 1(중앙대학교) 김다혜 /Dahye Kim 2(중앙대학교)
The Social Equity of Urban Parks in High-Density Urban Areas: A Case Study in the Core Area of Beijing	Kaiqi Zhang, Mingze Chen
GIS-based multi-criteria analysis for sustainable urban green spaces planning in emerging towns of Ethiopia: the case of Sululta town	Eshetu Gelan
Green space accessibility helps buffer declined mental health during the COVID-19 pandemic: evidence from big data in the United Kingdom	Kwan Ok Lee, Ke Michael Mai & Sounil Park
Multi-method analysis of urban green space accessibility: Influences of land use, greenery types, and individual characteristics factors	Kaiqi Zhang, Mingze Chen
Demand for parks and perceived accessibility as key determinants of urban park use behavior	Jingyuan Zhang, Puay Yok Tan
MIT Terrascope 2024 – Economic Feasibility (Terrascope Webpage) 검색일자. 2025.06.21	<a href="https://terrascope2024.mit.edu/?page_id=158">https://terrascope2024.mit.edu/?page_id=158</a>
Impacts of perceived safety and beauty of park environments on time spent in parks: Examining the potential of street view imagery and phone-based GPS data	Hanlin Zhou, Jue Wang, Kathi Wilson
뉴스 1, “서울 35도 넘긴 날 9→58일로 늘어”...‘열섬효과’로 대도시 폭염 위험, 검색일자:25.06.21	<a href="https://www.donga.com/news/Society/article/all/20240628/125670797/1">https://www.donga.com/news/Society/article/all/20240628/125670797/1</a>