

서울시 구별 활동지수 산출을 통한 지역 활성화 방안 제언



범주형자료분석팀 | 김보현 송승현 최용원 김동희 오주원

목차

01 주제 선정 배경

- 서울시 지역 불균형
- 활동지수의 필요성

- 안심구역
- 상호제공
- 외부데이터

02 데이터 설명

03 활동지수 산출

- 지수 분류
- 지수 산출 방식
- 최종 지수

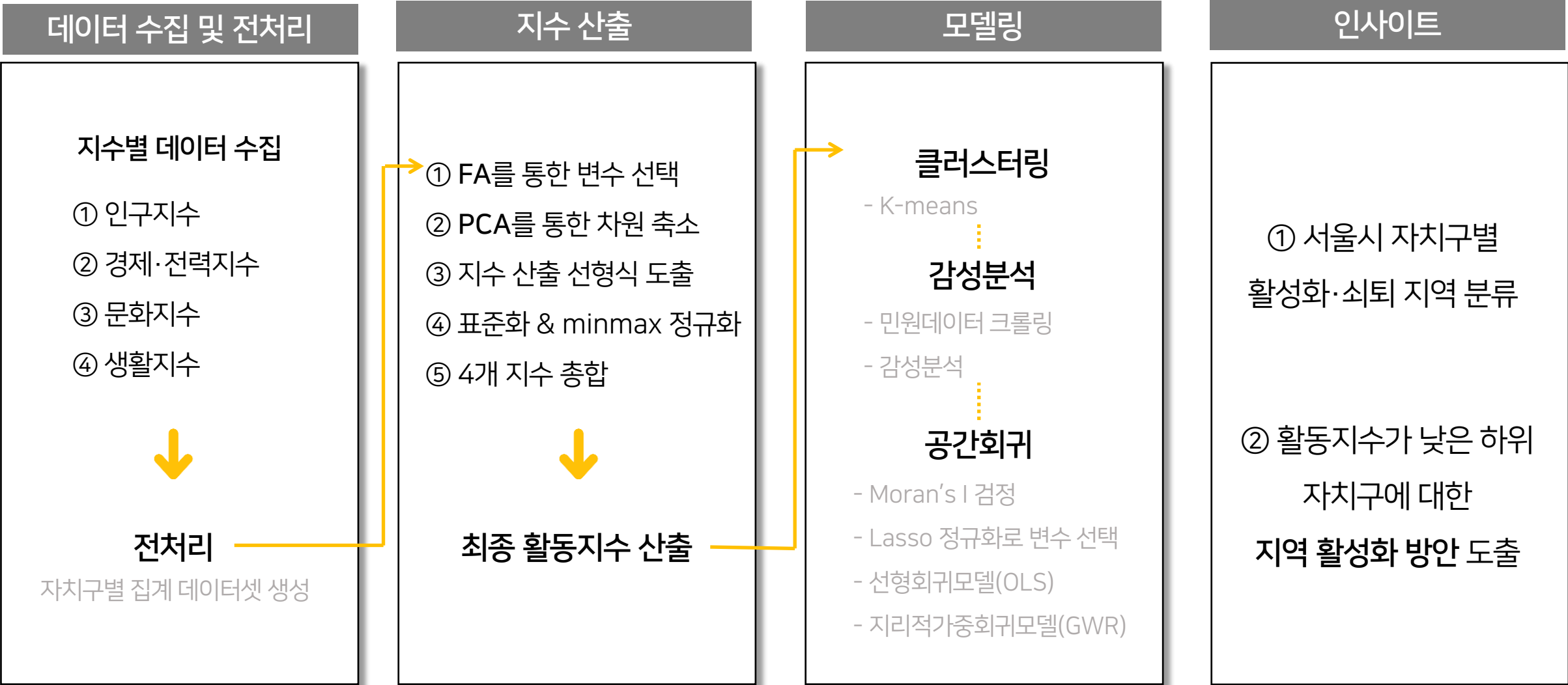
- 클러스터링
- 감성분석
- 공간회귀

04 모델링

05 결론

- 인사이트
- 의의 및 유의할 점
- 기대효과

분석 흐름 요약



01 주제 선정 배경

서울시 지역 불균형 | 활동지수의 필요성

서울시 지역 불균형

지역 간 격차

사회-경제적 수준의 공간적 편차

지역 불균형

사회적인 기회나 자원 등이 지역 간에 골고루
분포하지 못해 큰 차이가 발생하는 상태

서울 및 수도권 중심의
경제성장 정책으로 인한
지역 간 불균형

거주자의 상대적 박탈감과
사회 · 정치적 문제 유발
국가 전체의 경제성장을 억제

지역 불균형 해소를 위한
정부 개입 필요

이경영 외, <서울시 강남-강북권 간 지역격차 및 주거특성이
주민의 사회적 지위인식에 미치는 영향>

서울시 지역 불균형

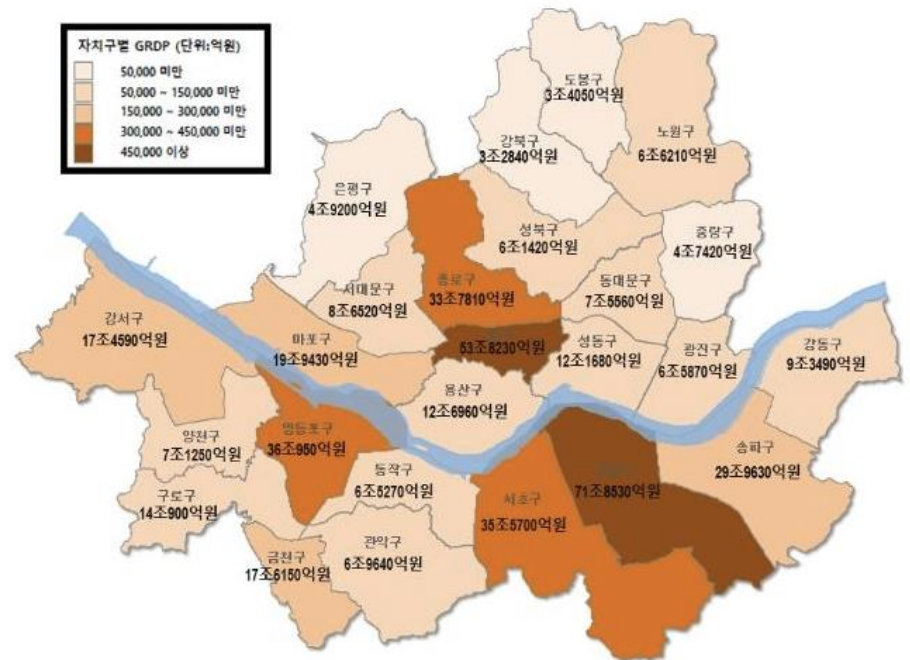
서울시 지역박탈지수

| 자치구 | 지수 |
|-----|-------|
| 서초구 | -8.25 |
| 강남구 | -7.12 |
| 송파구 | -5.07 |
| 중랑구 | 5.73 |
| 금천구 | 6.11 |
| 강북구 | 6.51 |

박탈 정도가 심한 상위 3개, 양호한 하위 3개 자치구

서울시 내 자치구 별로 박탈 정도에 **큰 차이**를 보임

→ 낙후된 지역에 민간 투자 저하 및 자원의 효율적인 이용 불가
지역 주민들 간 갈등으로 주민 통합 저해



서울특별시, <2019년 서울시 지역내 총생산(GRDP)>

경제력에서도 자치구 간 **엄청난 차이**를 보임

→ 서울시 경제력 1위 자치구는 강남구이며
이는 강북구의 22배에 해당

서울시 지역 불균형

서울시 지역균형발전계획

지역간 격차 해소와 지역별 특성 있는 발전을 위한 종합적인 계획

산업·일자리 기반 조성, 생활 SOC 확충, 지역 간 교통 격차 해소, 주거환경 개선, 교육 기회 등 사회 전반 전략

연합뉴스 2022.12.01

▶ 권역별 전략사업

- 서북권 - 역세권 중심 전략산업 육성
- 서남권 - 저이용 부지를 활용한 복합개발, 창업허브 조성
- 동북권 - 복합문화시설 건립, 미래산업 클러스터 조성
- 동남권 - 첨단산업 육성, 미래 성장동력 확보
- 도심권 - 녹지공간 확보 및 녹지생태도심 추진

서울시, 균형발전 5개년 계획 추진... 권역별 전략사업 육성

서울신문 2022.12.02

지역격차 해소·균형발전... 서울시, 5개년 계획 발표

서울시는 1일 지역 격차 해소와 균형발전을 위한 5개년 계획을 발표했다. 서울을 5개 권역으로 나눠 지역별 특성을 살려 전략사업을 육성한다는 계획이다.

서울시 지역균형발전계획은 서울시 조례에 따라 5년마다 수립하게 돼 있으며 이번 계획은 올해부터 2026년까지다.

...(후략)

활동지수의 필요성



기존 지표에 대한 문제

- ① 신규 사업의 추진 여부에 초점을 맞춰
구체적이고 객관적인 평가 항목 **부재**

ex) 도시 쇠퇴지수의 경우: 인구 감소, 사업체 수 감소,
생활환경 악화와 관련된 5개 지표만 활용

- ② 서울시 내의 지역균형을 세부적으로
판단하기에는 **부적합**

조운애, <도시 쇠퇴지수 개발에 관한 연구>



활동지수의 필요성

- ① 종합적인 지역균형발전계획의 성격을 고려하여
사회의 전 부문을 고려

ex) 인구, 경제·전력, 문화, 생활(교육, 보건 등)

- ② 자치구별 및 생활권별 균형 발전을 고려해
서울시 내 지역 균형 발전을 위한 **정책 제언**

최종 목표!

활용 데이터 - 한국전력공사 안심구역

전력과 경제활동 간의 관계

① 전력수요지수를 회귀식으로 모델링하여 경기동향지수를 분석한 결과,

실적 산업생산지수와 평균 오차범위가 약 0.41%로 **매우 유사**

$$\rightarrow \log(\text{월별 산업생산지수}) = \alpha + \text{탄력성} \times \log(\text{월별 전력수요지수}) + \varepsilon_t$$

② 전력수요지수는 지역별, 산업별 통계가 가능하므로 산업별 및 지역별 경기동향지표 개발 가능

이에 따라 **전력수요와 경기는 밀접한 관련성을 가짐**

지역경제단위의 경기동향자료 파악이 열악한 경우 전력수요량으로 대체하여 조기 예측 가능

활용 데이터 - 한국전력공사 안심구역

[한국전력공사] 관공/상업/주거용도 사용고객 시간대별 전력사용량

| 고객번호 | 용도별 | 연도 | 월 | 일 | 시간 (0015~2400) | 전력사용량 |
|---------|---------|---------|---------|--------|-------------------|--------|
| VARCHAR | VARCHAR | VARCHAR | VARCHAR | NUMBER | NUMBER | NUMBER |
| 1 | 관공용 | 2018 | 3 | 15 | 0015 | 0.** |

용도별로 구분되어 있는 전력수요량 안심구역 데이터를 활용하여
지역별 활동지수 산출 시 기존의 경제 관련 지수에 전력 관련 변수를 추가

→ 세부적인 지역경제 파악이 가능해짐

활용 데이터 - 상호제공 미개방데이터

| 번호 | 데이터명 | 제공기관명 | 안심구역 | 활용지수 |
|----|---|-----------|-------------|-------|
| 1 | 전국 직종별 경제활동 인구 현황 | 국민연금 | 전라북도·국민연금공단 | 인구 |
| 2 | 전국 직업분류별 평균소득 현황 | 국민연금 | 전라북도·국민연금공단 | 경제·전력 |
| 3 | 전국 사업장 업종별 규모 현황 | 국민연금 | 전라북도·국민연금공단 | 경제·전력 |
| 4 | 자동차 관련 정보, 출퇴근 정보, 부동산 소유 정보 | KCB | K-DATA | 경제·전력 |
| 5 | 지역별 가맹점 정보, 매출 및 이용고객 정보, 배달매출 및 이용고객 정보 | 신한카드 | K-DATA | 경제·전력 |
| 8 | 지역 인구특성 데이터 | NICE지니데이터 | K-DATA | 경제·전력 |

활용 데이터 - 외부데이터

인구지수

- 서울시 행정구역 (동별) 통계
- 서울시 인구이동률(구별) 통계
- 서울시 1인가구(연령별) 통계

⋮

경제·전력지수

- 서울시 전력 사용량 (용도별) 통계
- 신재생 에너지 계약현황

문화지수

- 서울시 공원 (1인당 공원면적) 통계
- 서울시 공공도서관 (구별) 통계
- 서울시 박물관 현황 통계

⋮

생활지수

- 서울시 녹지현황 통계
- 서울시 의료기관 (구별) 통계
- 서울시 건강보험 급여 (구별) 통계

⋮

자세한 데이터 활용 목록은 부록 참고

상호제공 미개방데이터와 외부데이터를 활용해 **지역별 활동지수 산출**

지수 분류

지수 관련 사전연구 검토

① NABIS 균형발전지표

- 인구, 경제, 주거, 교통, 산업, 교육, 환경 등

② 농촌경제연구원 2020 지역발전지수

- 생활서비스, 지역 경제력, 삶의 여유공간 등

③ 도시재생종합정보체계 도시재생 분석지표

- 노후건축물비율, 노령화지수, 신규주택비율 등



도시의 활성화 정도를 파악할 지수 산출

인구지수

인구, 가구, 주택 등을
종합적으로 고려

경제·전력지수

전력 수요와 경제 발전의
관계성을 중점적으로 통합

문화지수

문화시설 및 대중교통,
녹지면적 등 고려

생활지수

교육, 건강 등 생활 전반

03 활동지수 산출

지수 분류 | 지수 산출 방식 | 최종 지수

지수 산출 방식

자치구별 인구지수 상관관계

| | | | | | | | | | | | | | | | | | | |
|-------------|-------|-------|-------|--------|-------|-------|-------|--------|-------|--------|---------|-------|--------|--------|-------|-------|-------|---------|
| 전체인구수 | 1 | 0.99 | 0.97 | 0.99 | 0.66 | 0.66 | 0.32 | 0.31 | 0.62 | 0.97 | 0.97 | 0.25 | 0.017 | 0.05 | -0.62 | 1 | 0.67 | 0.075 |
| 경제활동인구수 | 0.99 | 1 | 0.99 | 0.98 | 0.66 | 0.65 | 0.32 | 0.32 | 0.6 | 0.95 | 0.99 | 0.16 | 0.073 | 0.14 | -0.61 | 1 | 0.73 | 0.059 |
| 취업인원수 | 0.97 | 0.99 | 1 | 0.95 | 0.65 | 0.65 | 0.3 | 0.32 | 0.52 | 0.91 | 0.97 | 0.12 | 0.16 | 0.23 | -0.64 | 0.98 | 0.74 | 0.12 |
| 실업인원수 | 0.99 | 0.98 | 0.95 | 1 | 0.63 | 0.63 | 0.34 | 0.33 | 0.69 | 0.96 | 0.96 | 0.25 | -0.048 | -0.05 | -0.55 | 0.98 | 0.65 | 0.033 |
| 면적 | 0.66 | 0.66 | 0.65 | 0.63 | 1 | 1 | -0.44 | -0.45 | 0.34 | 0.71 | 0.62 | 0.4 | 0.044 | 0.2 | -0.5 | 0.67 | 0.4 | 0.33 |
| 구성비 | 0.66 | 0.65 | 0.65 | 0.63 | 1 | 1 | -0.45 | -0.45 | 0.34 | 0.71 | 0.62 | 0.4 | 0.045 | 0.2 | -0.5 | 0.66 | 0.4 | 0.32 |
| 면적대비전체인구수 | 0.32 | 0.32 | 0.3 | 0.34 | -0.44 | -0.45 | 1 | 0.99 | 0.25 | 0.23 | 0.33 | -0.2 | -0.08 | -0.2 | -0.2 | 0.32 | 0.24 | -0.25 |
| 면적대비경제활동인구수 | 0.31 | 0.32 | 0.32 | 0.33 | -0.45 | -0.45 | 0.99 | 1 | 0.23 | 0.2 | 0.34 | -0.29 | -0.019 | -0.1 | -0.2 | 0.32 | 0.31 | -0.26 |
| 국민기초생활보장수급자 | 0.62 | 0.6 | 0.52 | 0.69 | 0.34 | 0.34 | 0.25 | 0.23 | 1 | 0.62 | 0.66 | 0.073 | -0.29 | -0.33 | 0.02 | 0.59 | 0.53 | -0.19 |
| 주택수 | 0.97 | 0.95 | 0.91 | 0.96 | 0.71 | 0.71 | 0.23 | 0.2 | 0.62 | 1 | 0.91 | 0.45 | -0.11 | -0.074 | -0.65 | 0.96 | 0.53 | 0.15 |
| 일반가구수 | 0.97 | 0.99 | 0.97 | 0.96 | 0.62 | 0.62 | 0.33 | 0.34 | 0.66 | 0.91 | 1 | 0.045 | 0.077 | 0.18 | -0.49 | 0.97 | 0.83 | -0.0046 |
| 주택보급률 | 0.25 | 0.16 | 0.12 | 0.25 | 0.4 | 0.4 | -0.2 | -0.29 | 0.073 | 0.45 | 0.045 | 1 | -0.4 | -0.5 | -0.46 | 0.22 | -0.45 | 0.31 |
| 충전입률 | 0.017 | 0.073 | 0.16 | -0.048 | 0.044 | 0.045 | -0.08 | -0.019 | -0.29 | -0.11 | 0.077 | -0.4 | 1 | 0.74 | -0.16 | 0.033 | 0.28 | 0.08 |
| 충전출률 | 0.05 | 0.14 | 0.23 | -0.05 | 0.2 | 0.2 | -0.2 | -0.1 | -0.33 | -0.074 | 0.18 | -0.5 | 0.74 | 1 | -0.12 | 0.1 | 0.5 | 0.14 |
| 노령화지수 | -0.62 | -0.61 | -0.64 | -0.55 | -0.5 | -0.5 | -0.2 | -0.2 | 0.02 | -0.65 | -0.49 | -0.46 | -0.16 | -0.12 | 1 | -0.64 | -0.12 | -0.31 |
| 통근-통학인구수 | 1 | 1 | 0.98 | 0.98 | 0.67 | 0.66 | 0.32 | 0.32 | 0.59 | 0.96 | 0.97 | 0.22 | 0.033 | 0.1 | -0.64 | 1 | 0.69 | 0.081 |
| 1인가구수 | 0.67 | 0.73 | 0.74 | 0.65 | 0.4 | 0.4 | 0.24 | 0.31 | 0.53 | 0.53 | 0.83 | -0.45 | 0.28 | 0.5 | -0.12 | 0.69 | 1 | -0.14 |
| 자녀거주인구수 | 0.075 | 0.059 | 0.12 | 0.033 | 0.33 | 0.32 | -0.25 | -0.26 | -0.19 | 0.15 | -0.0046 | 0.31 | 0.08 | 0.14 | -0.31 | 0.081 | -0.14 | 1 |

인구지수

자치구별 문화지수 상관관계

| | | | | | | | | | | | | |
|--------------|--------|--------|-------|--------|-------|--------|-------|--------|--------|-------|-------|--------|
| 방문자수 | 1 | 1 | 0.25 | -0.059 | 0.054 | 0.32 | 0.17 | 0.032 | 0.45 | 0.45 | 0.32 | 0.14 |
| 방문자비용 | 1 | 1 | 0.25 | -0.062 | 0.053 | 0.31 | 0.17 | 0.031 | 0.45 | 0.45 | 0.32 | 0.14 |
| 문화시설수합계 | 0.25 | 0.25 | 1 | -0.17 | 0.82 | 0.48 | 0.51 | 0.13 | -0.33 | -0.22 | 0.45 | 0.16 |
| 문예회관수 | -0.059 | -0.062 | -0.17 | 1 | -0.4 | 0.13 | -0.29 | 0.052 | 0.021 | 0.035 | -0.41 | -0.085 |
| 종합사회복지관수 | 0.054 | 0.053 | 0.82 | -0.4 | 1 | 0.14 | 0.29 | 0.25 | -0.26 | -0.25 | 0.4 | 0.25 |
| 구민(다목적)체육시설수 | 0.32 | 0.31 | 0.48 | 0.13 | 0.14 | 1 | -0.22 | -0.18 | 0.088 | 0.11 | 0.23 | -0.047 |
| 청소년수련시설수 | 0.17 | 0.17 | 0.51 | -0.29 | 0.29 | -0.22 | 1 | 0.077 | -0.42 | -0.23 | 0.29 | 0.04 |
| 공원면적 | 0.032 | 0.031 | 0.13 | 0.052 | 0.25 | -0.18 | 0.077 | 1 | -0.086 | 0.14 | 0.1 | 0.3 |
| 문화활동횟수 | 0.45 | 0.45 | -0.33 | 0.021 | -0.26 | 0.088 | -0.42 | -0.086 | 1 | 0.55 | 0.13 | 0.12 |
| 박물관수 | 0.45 | 0.45 | -0.22 | 0.035 | -0.25 | 0.11 | -0.23 | 0.14 | 0.55 | 1 | 0.11 | 0.21 |
| 공공도서관개수 | 0.32 | 0.32 | 0.45 | -0.41 | 0.4 | 0.23 | 0.29 | 0.1 | 0.13 | 0.11 | 1 | 0.5 |
| 공공도서관좌석수 | 0.14 | 0.14 | 0.16 | -0.085 | 0.25 | -0.047 | 0.04 | 0.3 | 0.12 | 0.21 | 0.5 | 1 |

문화지수

각 지수별 변수 간의 높은 상관관계 확인



PCA, FA 등 차원 축소 방법 고려해 지수 산출

03 활동지수 산출

지수 분류 | 지수 산출 방식 | 최종 지수

지수 산출 방식



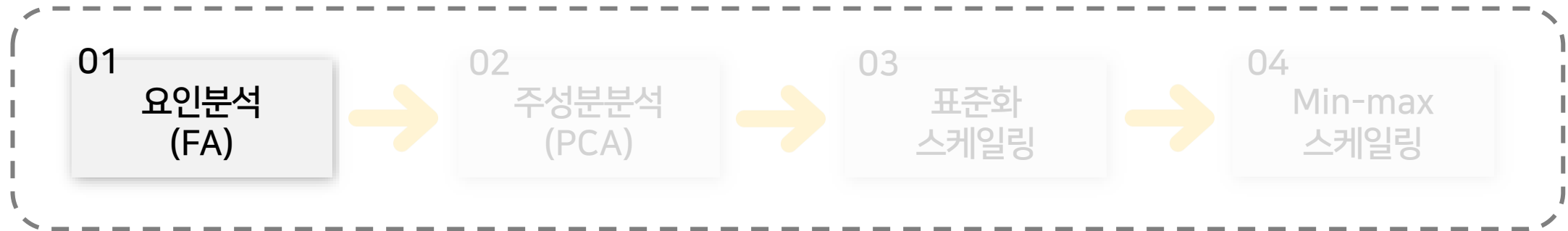
최종 지수 산출 흐름

- | | |
|----|---|
| 01 | 요인분석(FA)을 통해 지수 산출에 활용될 변수 선택 |
| 02 | 주성분분석(PCA)에서 PC1의 loading값을 각 변수의 가중치로 채택 |
| 03 | 표준화 스케일링을 통해 변수 간 범위 재조정 |
| 04 | Min-max 스케일링으로 지수 산출 |
| 05 | 4개의 지수를 총합해, 최종 지수 산출 |

03 활동지수 산출

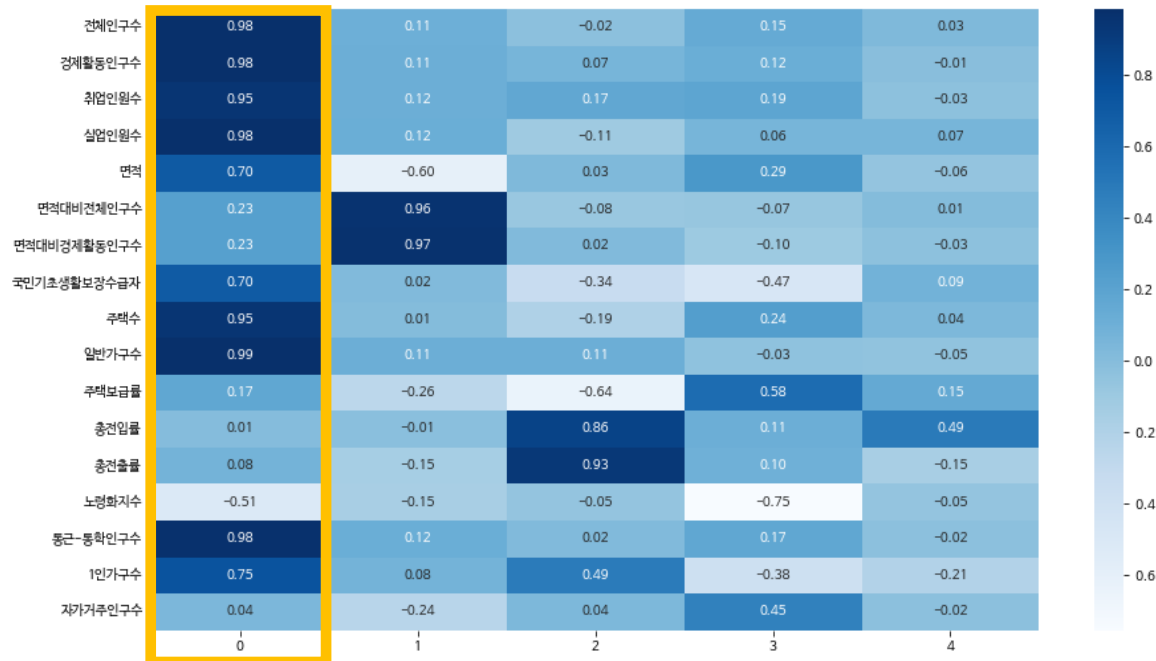
지수 분류 | 지수 산출 방식 | 최종 지수

지수 산출 방식



* 요인분석(FA): 변수들 간의 상호 관련성을 분석해, 잠재요인을 찾아 각 변수의 영향력을 추출하는 기법

ex) 인구지수 FA 결과



요인분석 결과,
factor0에서 loading값이 높은 변수들만 채택

- 인구지수 산출에 사용된 최종 변수
전체인구수, 경제활동인구수, 취업인원수, 실업인원수,
주택수, 일반가구수, 통근-통학인구수

* Loading값 0.9 이상

03 활동지수 산출

지수 분류 | 지수 산출 방식 | 최종 지수

지수 산출 방식



* 주성분분석(PCA): 변수 간 상관관계를 활용해 기존의 분산을 최대한 보존하는 차원축소 기법

ex) 인구지수 PCA 결과

| | PC1 | PC2 | PC3 | PC4 |
|----------|----------|-----------|-----------|-----------|
| 전체인구수 | 0.382429 | -0.111434 | 0.017960 | 0.119872 |
| 경제활동인구수 | 0.382268 | 0.174675 | -0.054660 | -0.100809 |
| 취업인원수 | 0.375407 | 0.413937 | -0.628735 | 0.301955 |
| 실업인원수 | 0.378288 | -0.213166 | 0.509801 | 0.683233 |
| 주택수 | 0.369494 | -0.737116 | -0.206354 | -0.378698 |
| 일반가구수 | 0.375364 | 0.443408 | 0.521783 | -0.507880 |
| 통근-통학인구수 | 0.382313 | 0.018326 | -0.163234 | -0.127003 |

인구지수 : 0.9727

PC1의 PVE가 충분히 높을 경우

PC1의 loading을 각 변수의 가중치로 활용 가능

* loading: PC에 대한 각 변수의 영향력

*소수 둘째자리에서 반올림

$$\begin{aligned} \text{인구지수} = & 0.38 * \text{전체인구수} + 0.38 * \text{경제활동인구수} + 0.38 * \text{취업인원수} + 0.38 * \text{실업인원수} \\ & + 0.37 * \text{주택수} + 0.38 * \text{일반가구수} + 0.38 * \text{통근통학인구수} \end{aligned}$$

03 활동지수 산출

지수 분류 | 지수 산출 방식 | 최종 지수

지수 산출 방식



- Standard scaling

$$x_{scaled} = \frac{x - \mu}{\delta}$$

- Min-max scaling

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

- 최종 활동 지수

인구지수 + 경제·전력지수
+ 문화지수 + 생활지수

각 변수를 Standard scaling 후,
선형식으로 지수 산출식 도출

Min-max scaling으로 각각의 지수 산출
- 값의 범위를 0~1 사이로 변경해, 자치구별 지수를 쉽게 비교 가능

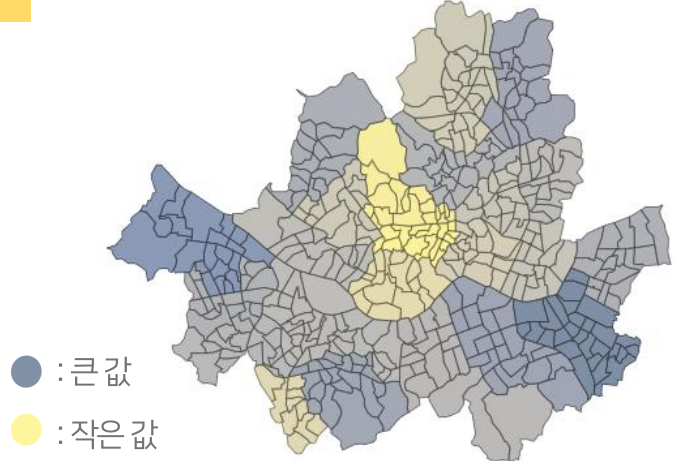



4개의 지수를 총합해 **최종 활동 지수** 산출


03 활동지수 산출

지수 분류 | 지수 산출 방식 | 최종 지수

① 인구지수

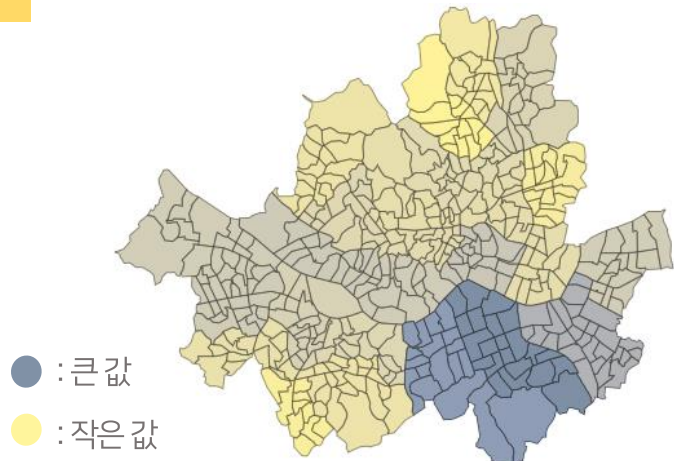


 지수 **높은** 순: 송파구, 강서구, 강남구, 관악구, 노원구

 지수 **낮은** 순: 중구, 종로구, 용산구, 금천구, 성동구

| 자치구 | 전체인구수 | 경제활동 인구수 | 취업인원수 | 실업인원수 | 주택수 | 일반가구수 | 통근- 통학인구수 | 산출 | 지수 |
|-----|----------|-------------|---------|---------|--------|--------|--------------|-------|-------|
| 송파구 | 19773406 | 12444230 | 6350962 | 3002712 | 211578 | 255766 | 376433 | 5.79 | 1 |
| 강서구 | 17226825 | 10941217 | 5601412 | 2792936 | 195383 | 244097 | 326875 | 4.43 | 0.877 |
| ... | | | | ... | | | | | ... |
| 종로구 | 4314837 | 2689486 | 1110354 | 657840 | 45859 | 63414 | 82692 | -4.97 | 0.035 |
| 중구 | 3668958 | 2310609 | 1079236 | 531574 | 42234 | 55093 | 68167 | -5.37 | 0 |

② 경제·전력지수

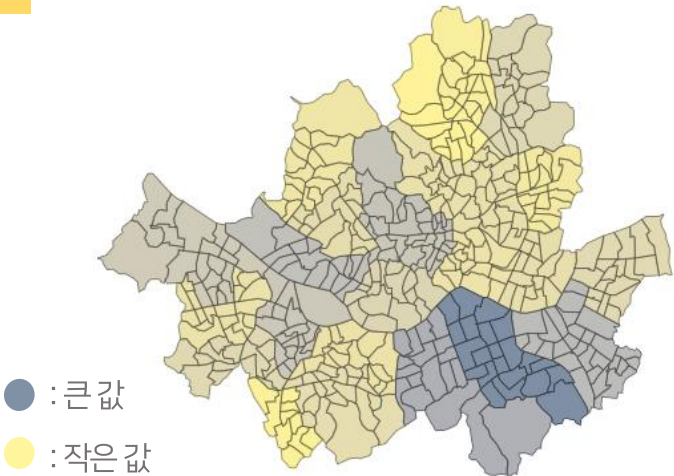




지수 높은 순: 강남구, 서초구, 송파구, 성동구, 마포구

지수 낮은 순: 강북구, 금천구, 중랑구, 도봉구, 은평구

| 자치구 | 총자산평균 평가금액 | 월평균카드 소비금액 | 평균부동산 매매가격 | 평균부동산 전세가격 | 평균연소득 | 변환 _주거용 | 변환 _관공용 | 변환 _상업용 | 산출 | 지수 |
|-----|---------------|---------------|---------------|---------------|-------|------------|------------|------------|-------|-------|
| 강남구 | 875777 | 12452 | 78303 | 36109 | 47106 | 90857 | 12843242 | 494338 | 7.53 | 1 |
| 서초구 | 830291 | 11507 | 75664 | 32722 | 47576 | 69790 | 20232107 | 270843 | 5.93 | 0.847 |
| ... | | | | ... | | | | | | ... |
| 금천구 | 339101 | 6212 | 16795 | 9961 | 33808 | 32799 | 147707 | 147707 | -2.57 | 0.031 |
| 강북구 | 314280 | 5788 | 17884 | 10435 | 33525 | 39887 | 58628 | 58628 | -2.90 | 0 |

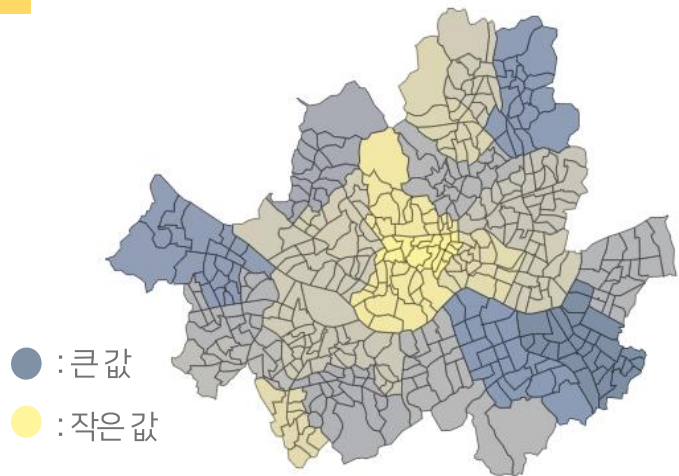
③ 문화지수



-  지수 **높은** 순: 강남구, 서초구, 송파구, 마포구, 종로구
-  지수 **낮은** 순: 강북구, 금천구, 중랑구, 도봉구, 은평구

| 자치구 | 문화시설이용자수 | 방문자수 | 방문자비율 | 대중교통이용승객수 | 녹지면적 | 산출 | 지수 |
|-----|----------|-----------|-------|-----------|---------|-------|-------|
| 강남구 | 1627101 | 404944228 | 11.2 | 261731880 | 2007604 | 6.57 | 1 |
| 서초구 | 570889 | 290336157 | 8 | 194508282 | 1447157 | 3.24 | 0.62 |
| ... | | | | ... | | | ... |
| 금천구 | 230644 | 56146190 | 1.6 | 65283441 | 388290 | -2.21 | 0.005 |
| 강북구 | 158654 | 59349220 | 1.6 | 88190101 | 138655 | -2.25 | 0 |

④ 생활지수



지수 높은 순: 송파구, 강남구, 노원구, 강서구, 은평구

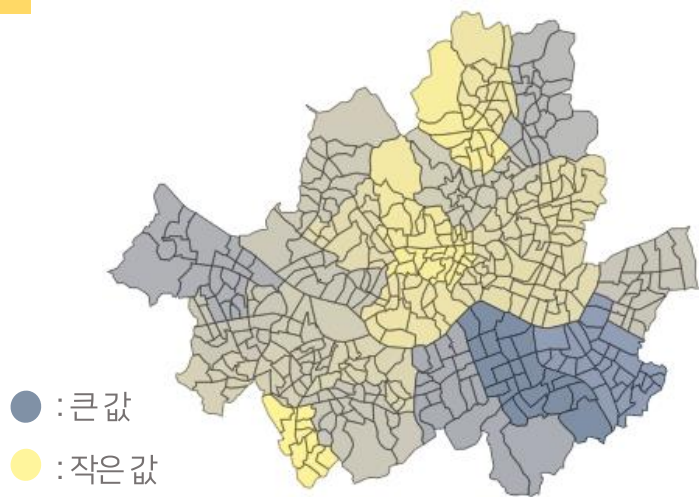
지수 낮은 순: 중구, 용산구, 종로구, 성동구, 금천구

| 자치구 | 공공체육시설개수 | 초등학교수 | 중학교수 | 고등학교수 | 건강보험 건수 | 주차장 개수 | 비상급수 시설개수 | 산출 | 지수 |
|-----|----------|-------|------|-------|----------|--------|-----------|-------|-------|
| 송파구 | 150 | 41 | 29 | 20 | 15194128 | 22566 | 1349 | 4.82 | 1 |
| 강남구 | 292 | 33 | 24 | 22 | 11226155 | 20001 | 890 | 3.62 | 0.862 |
| ... | | | | ... | | | | | ... |
| 용산구 | 40 | 15 | 9 | 10 | 4994636 | 7437 | 176 | -3.14 | 0.079 |
| 중구 | 32 | 12 | 8 | 11 | 3083967 | 4263 | 200 | -3.83 | 0 |

03 활동지수 산출

지수 분류 | 지수 산출 방식 | 최종 지수

최종 지수



지수 높은 순: 강남구, 송파구, 서초구, 강서구, 노원구



지수 낮은 순: 금천구, 중구, 강북구, 종로구, 도봉구

| 자치구 | 인구min_max | 경제min_max | 문화min_max | 생활min_max | 최종min_max |
|-----|-----------|-----------|-----------|-----------|-----------|
| 강남구 | 0.784821 | 1 | 0.862356 | 1 | 3.610777 |
| 송파구 | 1 | 0.590060 | 1 | 0.498290 | 3.088350 |
| ... | | | | ... | |
| 중구 | 0 | 0.194468 | 0 | 0.368785 | 0.563252 |
| 금천구 | 0.214586 | 0.031509 | 0.194928 | 0.004824 | 0.445847 |

클러스터링

25개의 자치구를 대상으로 잠재된 지역 특성을 분석, 권역으로 구분하고자 **클러스터링** 시도

* 각 데이터의 유사성을 기반으로 다수의 그룹으로 나눠 그룹 간 특성을 확인하는 비지도학습

K-means

평균을 중심으로
클러스터 내 분산을 최소화

K-medoids

중심점을 이용해
K개의 클러스터로 분할

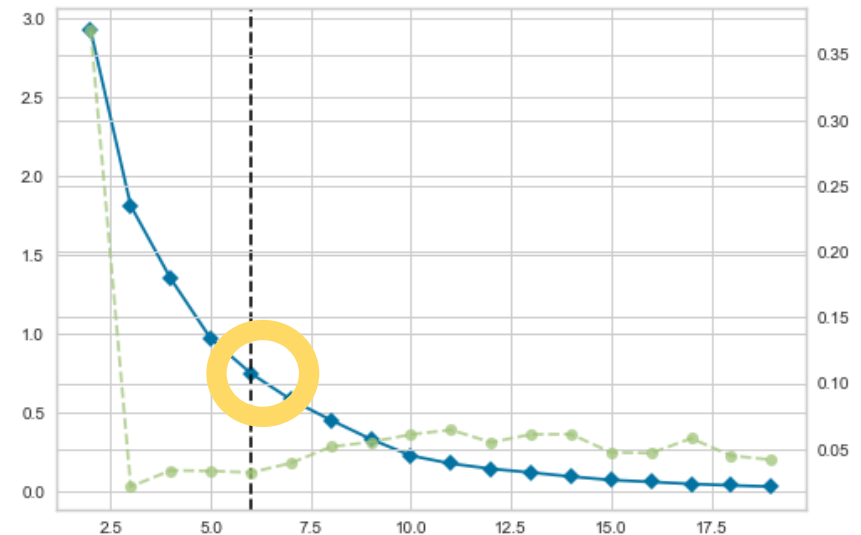
HC

계층적 클러스터링

DBSCAN

비계층적 클러스터링
밀도 기반

- ▶ 클러스터링 변수: 4개 지수
- ▶ 군집으로 잘 묶이지 않았던 HC, DBSCAN



Elbow point와 실루엣계수를 통해 최적의 클러스터 개수(K)
결정 실루엣계수가 가장 좋았던 **K-means(K=5)**로 결정

* Silhouette

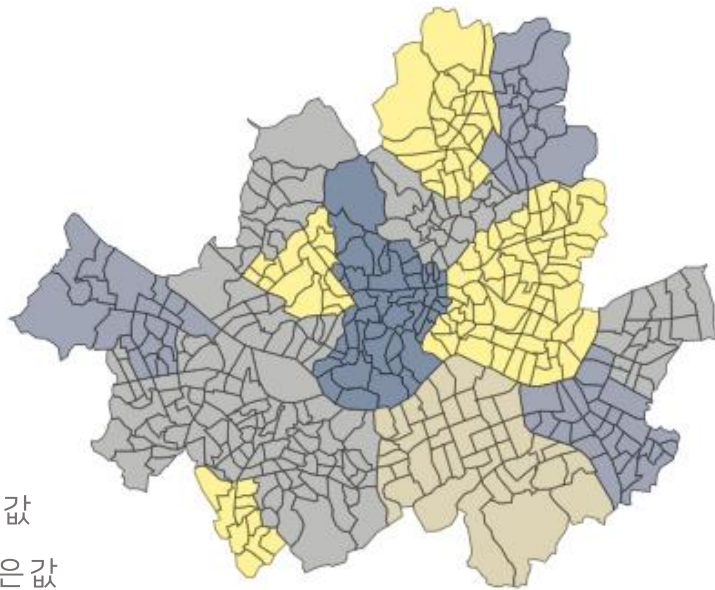
$$s(i) = \frac{b(i) - a(i)}{\max[a(i), b(i)]}$$

클러스터링

| 클러스터 | 인구지수 평균 | 경제·전력지수 평균 | 생활지수 평균 | 문화지수 평균 | 최종지수 평균 | 자치구 |
|------|----------|------------|----------|----------|----------|--|
| 1 | 0.354334 | 0.139656 | 0.297720 | 0.090589 | 0.882300 | 성동구, 광진구, 동대문구, 중랑구, 강북구, 도봉구, 서대문구, 금천구 |
| 2 | 0.635275 | 0.923564 | 0.698672 | 0.811629 | 3.069140 | 서초구, 강남구 |
| 3 | 0.566446 | 0.234503 | 0.495720 | 0.216083 | 1.512751 | 성북구, 은평구, 마포구, 양천구, 구로구, 영등포구, 동작구, 관악구, 강동구 |
| 4 | 0.865210 | 0.390847 | 0.892195 | 0.355255 | 2.503507 | 노원구, 강서구, 송파구 |
| 5 | 0.074734 | 0.231503 | 0.054625 | 0.343463 | 0.704324 | 종로구, 중구, 용산구 |

클러스터별 각 지수 평균에 유의미한 차이 존재
BUT 클러스터 당 자치구 개수가 일정하지 않음

클러스터링 | 결과



K-means 클러스터링 결과 시각화

- | | | |
|----|--------|---|
| 01 | 동(東)부권 | <ul style="list-style-type: none">● (서대문구, 금천구 제외) 동쪽 생활권● 경제, 문화 지수에서의 개선 필요 |
| 02 | 경제 특화권 | <ul style="list-style-type: none">● 경제 활동 중심의 광역 경제 생활권● 총자산금액, 평균 연소득 등 경제 전반에서 우수 |
| 03 | 서(西)부권 | <ul style="list-style-type: none">● (강동구, 성북구 제외) 서쪽 생활권● 생활, 경제, 문화 전반에서 균형적 |
| 04 | 거주 특화권 | <ul style="list-style-type: none">● 주요 거주 자치구● 경제, 문화 지수에서의 개선 필요 |
| 05 | 도심권 | <ul style="list-style-type: none">● 역사·문화 중심의 도심권● 인구, 생활 지수에서의 개선 필요 |

클러스터링 | 인사이트

앞선 클러스터링 결과와 2030 서울 생활권 발전 계획을 고려하였을 때 , 권역별 지수 향상을 위한 정책 제언 가능

| 클러스터 | 인구지수 평균 | 경제·전력지수 평균 | 생활지수 평균 | 문화지수 평균 | 최종지수 평균 | 자치구 |
|------|----------|------------|----------|----------|----------|--|
| 1 | 0.354334 | 0.139656 | 0.297720 | 0.090589 | 0.882300 | 성동구, 광진구, 동대문구, 중랑구, 강북구, 도봉구, 서대문구, 금천구 |

동(東)부권 고용 창출 및 문화복지시설, 공원 녹지 확충 필요

- 2030 서울플랜 (2014, 서울시)

- ✓ 창동상계 광역중심을 고용중심지로 육성
- ✓ 경춘선 폐선부지 공원화

| | | | | | | |
|---|----------|----------|----------|----------|----------|--------------|
| 5 | 0.074734 | 0.231503 | 0.054625 | 0.343463 | 0.704324 | 종로구, 중구, 용산구 |
|---|----------|----------|----------|----------|----------|--------------|

도심권 도심주거 확충, 노후주거지 개선, 교통 접근성 강화

- 2030 서울플랜 (2014, 서울시)

- ✓ 저층주거지, 철도역 배후지 등 노후주거지의 주거환경 개선
- ✓ 신안산선 등 도심과 연결되는 급행 간선 철도망 구축

감성분석 | 정의

텍스트에 나타난 사람들의 태도, 의견, 성향과 같은 주관적인 데이터를 분석하는 자연어 처리 기술

감성분석 진행방향

1. 데이터 수집(Data Collection)

- 서울시 자치구별로 특정 기간 동안의 전자 민원 데이터를 크롤링 방법을 통하여 수집

2. 텍스트 전처리(Text preprocessing)

- 감성 분석에 사용할 텍스트를 토큰화 등 전처리

3. 극성 탐지(Polarity Detection)

- 텍스트 안에 있는 긍정 · 부정적인 단어를 탐지, 이를 정량화 한 뒤 통계적 기법을 적용해 단어가 나타내는 점수의 총합이나 평균을 계산

감성분석 | 과정 및 활용 방안



서울시 자치구별 민원데이터 수집

감성분석에 사용할 텍스트 요소 분리

KNU 한국어 감성사전을 이용해 단어별 점수 계산

‘민원’임을 고려하여 부정적인 단어만 탐지해 자치구별 평균 점수 계산

활용 방안

- ② 공간회귀의 설명변수로 활용함으로써 최종지수에 미치는 영향 확인

감성분석 | 데이터수집

서대문구 | 새을

새을 전자민원창구

민원신청 및 조회

행정정보 공개

고객센터

공개 상담민원 조회

민원상담

2021-01-01

~

2021-12-31

제목

전체 (부서미지정 상)

답변여부(전체)

20개

조회

온라인 민원상담 (총216건)

민원상담 안내보기

| 번호 | 제목 | 작성자 | 작성일 | 답변여부(처리기한) | 조회 |
|--------|---|-----|------------|------------|----|
| 107633 | 홍은2동 일부 도로 배수트렌치 덮개 등 문제점 시정요청 | 김경호 | 2021-12-31 | 답변완료 | 16 |
| 107536 | 서대문구청 공영주차장 관리 시민 우회 | 김정철 | 2021-12-29 | 답변완료 | 23 |
| 107463 | 자가격리통지서 | 김광일 | 2021-12-27 | 답변완료 | 52 |
| 107401 | 전동킥보드 불법주차 단속 부탁드립니다. | 이다영 | 2021-12-24 | 답변완료 | 35 |
| 107390 | 상속확정등기 전 재산세 과세대상을 본인에서 배우자로 변경 가능한지요 ? | 박경하 | 2021-12-23 | 답변완료 | 9 |
| 107368 | 가로등 설치를 요청합니다. | 송○○ | 2021-12-23 | 답변완료 | 33 |
| 107333 | 마을버스 서대문 12번 노선연장 및 혼잡도 개선안에 대한 문의 | 김○○ | 2021-12-22 | 답변완료 | 32 |
| 107208 | 서대문구청 '주택 에너지 효율공사 지원' 보조금 지급신청서 제출 기한연장 요청 건 | 하○○ | 2021-12-20 | 답변완료 | 21 |



| 자치구 | 민원내용 |
|-----|-------------------------------|
| 강남구 | 주말 오전 시에 건물 공사를 하는게 말이 됩니까... |
| 강남구 | 자전거 공기 주입기 넣는 곳이 없어서 불편합니다... |
| 용산구 | 집 앞에 쓰레기 버리는 사람들 때문에 스트레스... |
| 금천구 | 해당 공간의 흡연자가 너무 많습니다... |
| 은평구 | 안심택배함 주변 상태 개선 요합니다... |
| 중랑구 | 오토바이 소음이 너무 시끄럽습니다... |
| 서초구 | 얼마 전 사무실 앞 인도 화단에 싱크홀이 발생... |
| 마포구 | 도로에 무단으로 주차하는 차량이 자꾸 발생해서... |

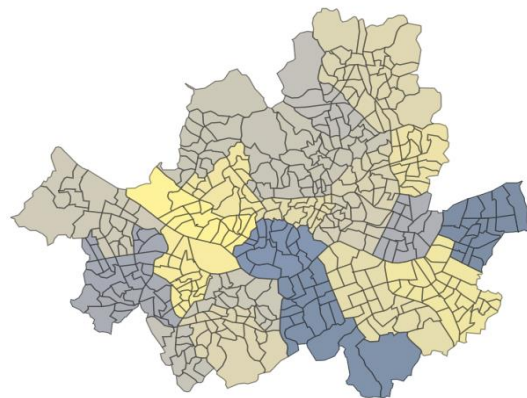
서울시 자치구 별 민원데이터를 '새을 전자민원창구'를 통해서 수집
총 9335개의 민원 데이터를 수집

감성분석 | 결과

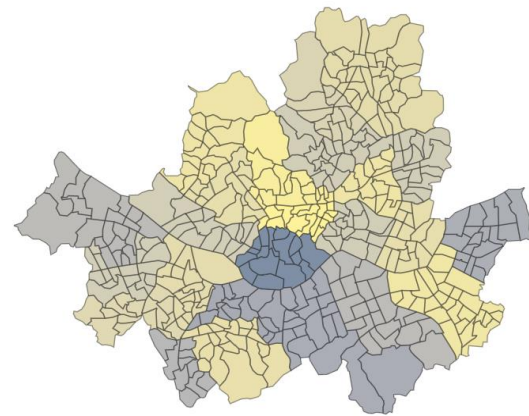
| 자치구 | avg_score | num |
|-----|-----------|-----|
| 용산구 | -2.858 | 579 |
| 동작구 | -3.497 | 398 |
| ... | | ... |
| 종로구 | -3.462 | 39 |
| 중구 | -3.7 | 20 |

*avg_score의 절댓값이 클수록 부정적임

평균 점수



민원 개수



● : 큰 값

● : 작은 값

부정적인 단어로 한정 지어 자치구별 평균 점수를 구한 결과, 자치구별로 감성분석 결과가 상이한 것을 확인

감성분석 점수가 가장 낮은 중구, 종로구가 생활지수도 가장 낮은 것으로 보아 근거로서 활용 가능

공간회귀 | 정의

지리학 제1법칙

“모든 것은 그 밖의 다른 모든 것과 관련되어 있지만, 인접해 있는 것들이 멀리 있는 것들보다는 더 높은 관련성을 보인다”

공간 효과

Moran's I 검정으로 자기상관 여부 확인

⋮

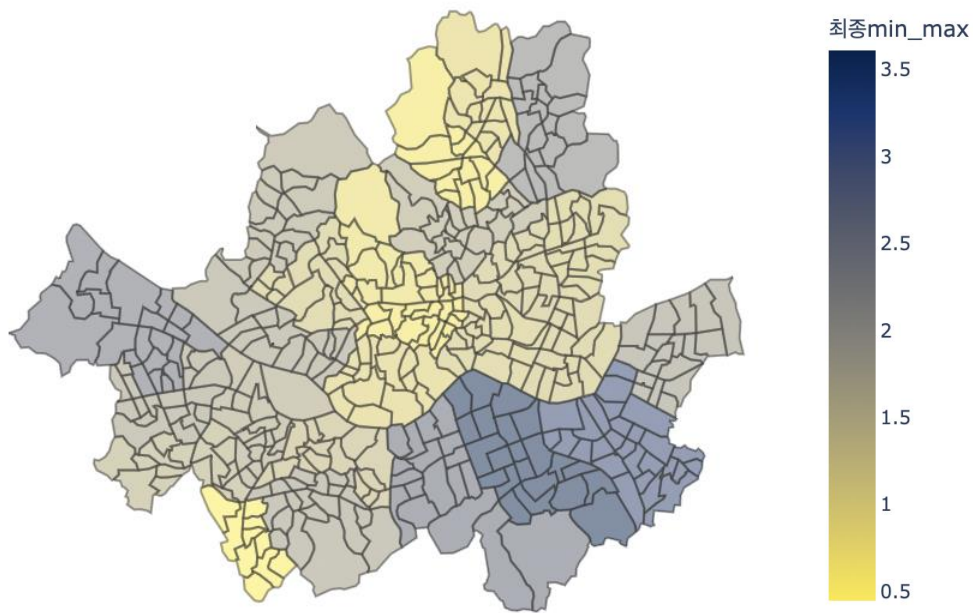
서울시 자치구의 공간적 자기상관여부 검정 결과,
Moran's I가 **0.57 ($p < .001$)**로 공간 간의 자기상관이 존재함을 확인

⋮

공간적 이질성으로 발생하는 이분산성을 고려하는
지역 간 거리 가중행렬을 이용하여 지역별 회귀계수 측정

소득수준 등 지역 활성화 지수에 사용한 자치구별 수집 변수는 **공간적 특성**을 지님 → 공간성을 고려한 **지리적 가중회귀모델** 사용

공간회귀 | 목표



자치구별로 최종지수가 상이한 것을 확인



서울특별시 홈페이지 2022. 12. 06

서울시 균형발전 5개년 계획 발표, 지역격차 해소 위한 발판 마련

- ① 산업·일자리 기반 조성
 - ② 균형발전 기반 생활SOC 확충
 - ③ 지역간 교통 격차 해소
 - ④ 주민 요구를 충족하는 주거환경 개선
 - ⑤ 균등한 교육 기회 제공
- 서울시 권역별, 자치구별로 지역격차 해소를 위한
5대 추진 전략 시행

서울시 균형발전 5개년 계획에 고려된 변수 추가 수집

- ① 산출한 최종 지수의 정당성 확인
- ② 기존에 지수에 사용하지 않았던 + 서울시 지역균형발전계획 도메인 지식 고려

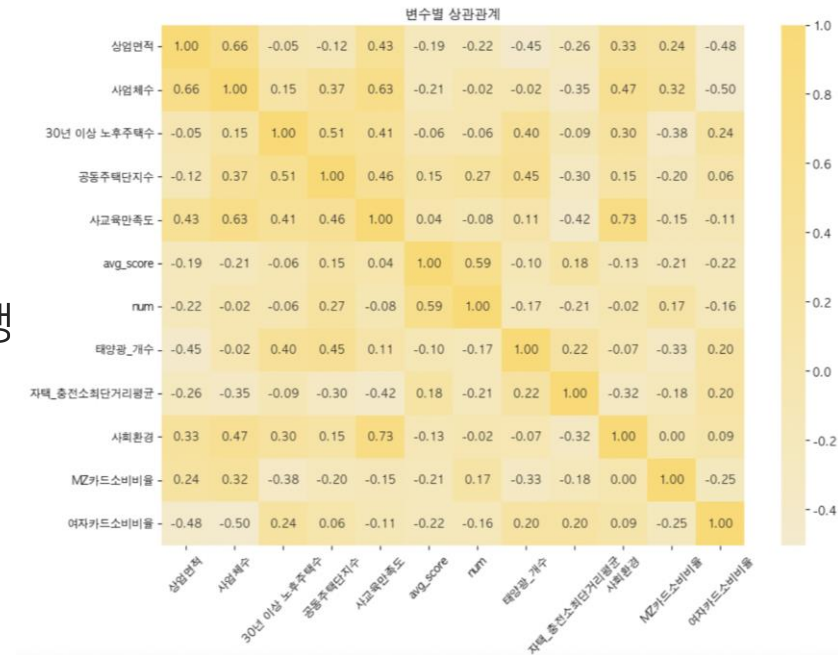


선택한 변수가 자치구별로 최종 지수 증가·감소에 미치는 영향 분석

공간회귀 | 변수 선택 과정

Lasso 정규화

- 제약조건을 통해 의미가 없는 변수들의 회귀계수를 0으로 만들어 변수 선택 가능
- 각 변수마다 값의 차이가 큼 → 종속변수인 최종 지수에 맞추어 각 변수 min-max scaling 진행
→ CV를 통해 최적의 λ 값 선택 : 0.01995
- Lasso 정규화로 선택된 변수에 대하여 VIF 측정 시 10 이상인 변수 없음 확인



▶ Lasso 정규화로 선택된 변수

주거환경 · 상권 측정 변수

- 상업면적, 사업체수, 30년 이상 노후주택 수, 공동주택단지수

민원 · 생활 관련 변수

- 사교육 만족도, 민원 점수, 민원 개수, 사회환경 만족도

소비 · 전력 관련 변수

- 태양광 개수, 자택 충전소 최단거리 평균, MZ카드소비비율, 여자카드소비비율

공간회귀 | 결과

선형회귀모델

- 관찰들 간의 독립성과 오차의 등분산성에 기초하여 공간적 변이의 측면 분석 어려움
- 공간적 위치에 따른 차이가 없다는 가정에 따라 변수의 효과가 모든 곳에서 일정하다고 전제

지리적 가중회귀모델

- 공간적 변이를 고려한 공간통계 분석 기법
 - ➔ 공간적 위치에 따라 다른 회귀계수로 모형화
 - 대역폭 설정에 따라 지리적 가중치 변화
 - ➔ Cross-Validation을 통한 최적의 대역폭 선택
- Lasso로 선택한 변수들로 진행했을 때 최적의 대역폭 10*

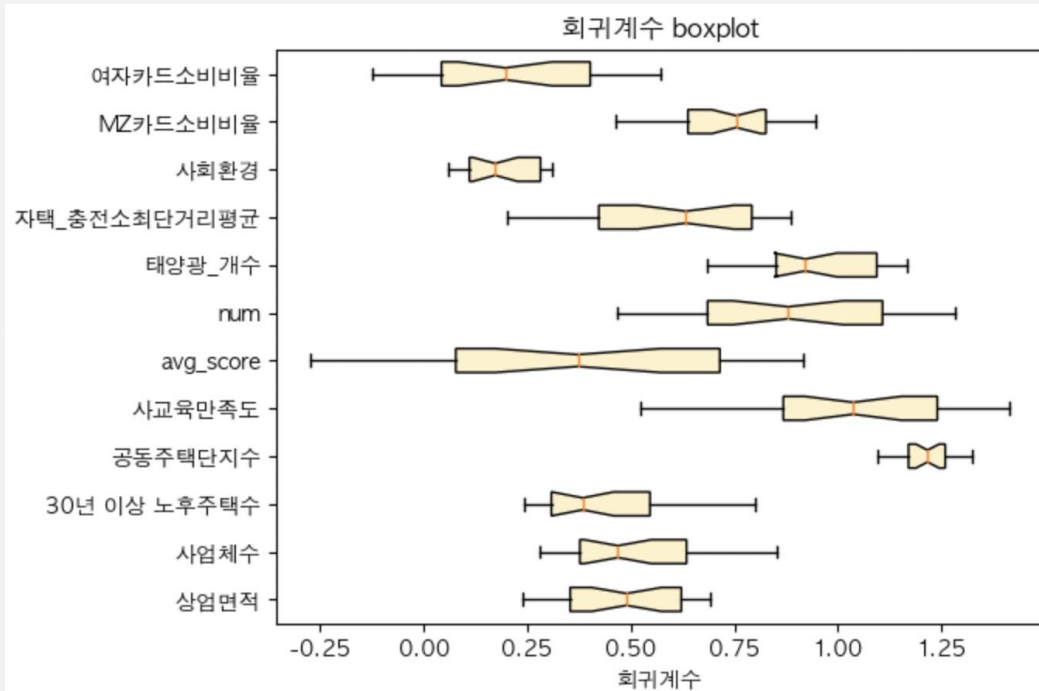
| | 선형회귀모델 (OLS) | 지리적가중회귀모델 (GWR) |
|-------------------|-----------------|--------------------|
| Adj. R-squared | 0.889 | 0.918 |
| AIC | 11.375 | -1.447 |
| BIC | 37.811 | 21.683 |

다양한 지표를 통해 선형회귀모델과 비교한 결과,
변수 간의 관계 고려 시

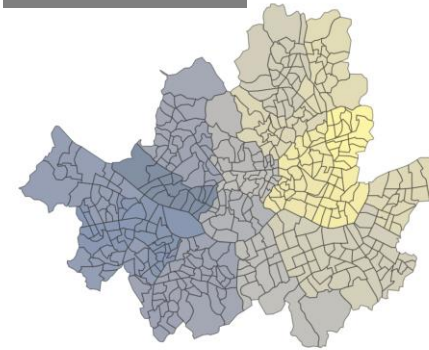
지리적 가중회귀모델에서 **공간 효과를 반영**한 것을 알 수 있음

공간회귀 | GWR 회귀계수 추정치

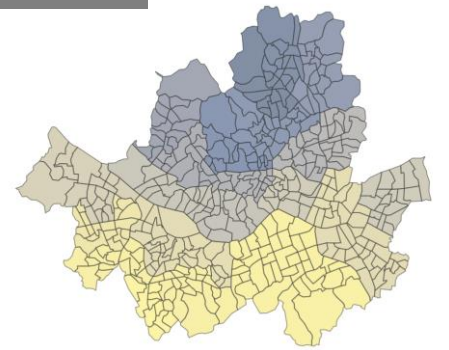
● : 회귀계수 값이 큼
● : 회귀계수 값이 작음



여성카드소비비율



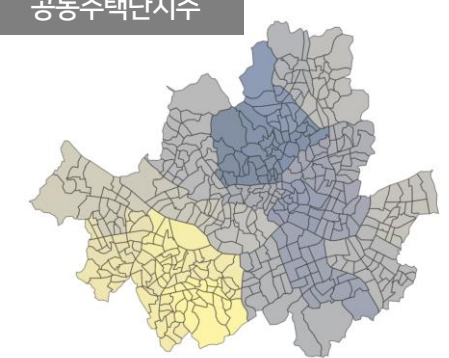
사업체수



민원 점수



공동주택단지수



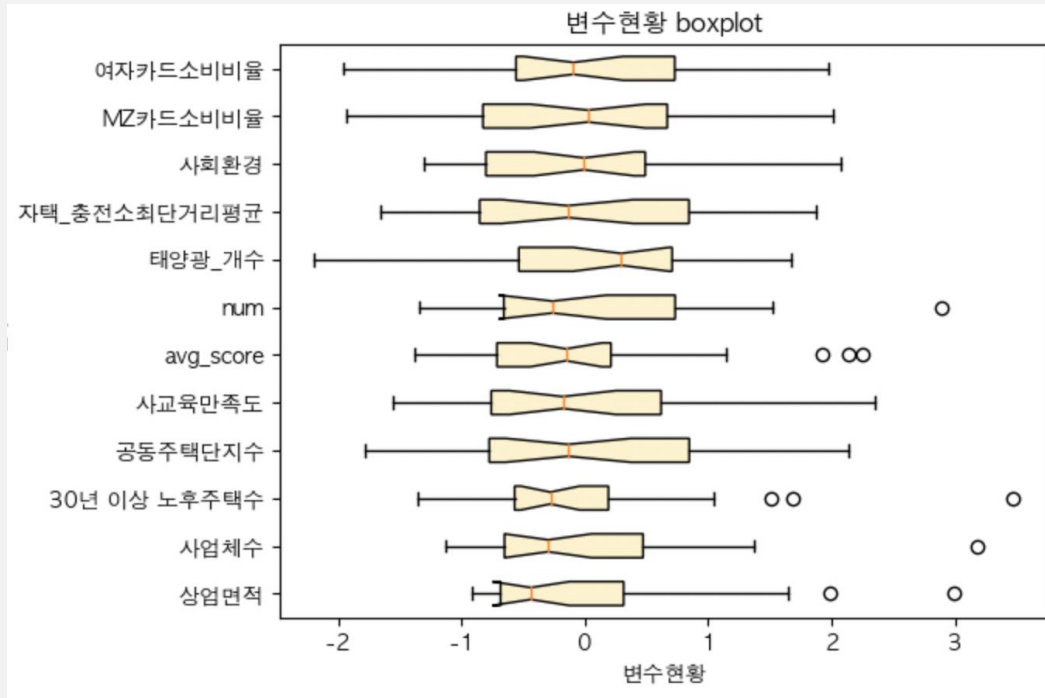
지리적 가중회귀모델 (GWR)을 적용한 결과, 각 변수의 회귀계수 추정치 해석 가능

회귀계수가 클수록 종속변수인 **최종지수**에 미치는 영향이 큰 것을 알 수 있음

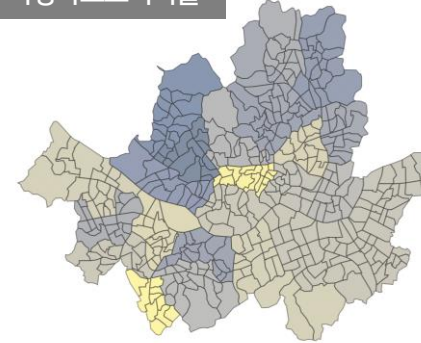


동일 변수 기존 자치구 현황과 비교하여 활성화 방안 도출

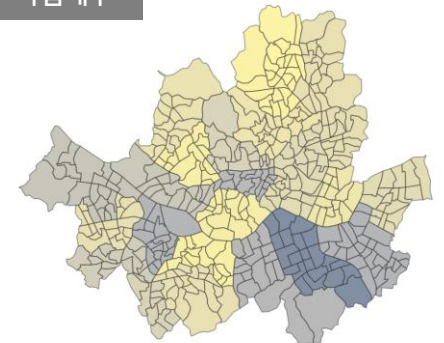
공간회귀 | 자치구별 변수 현황



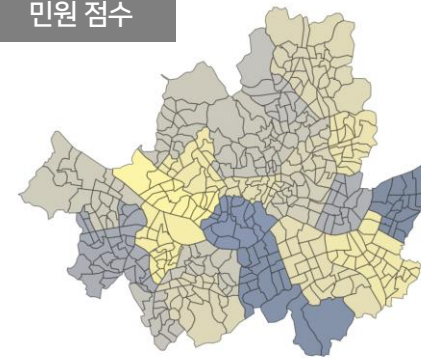
여성카드소비비율



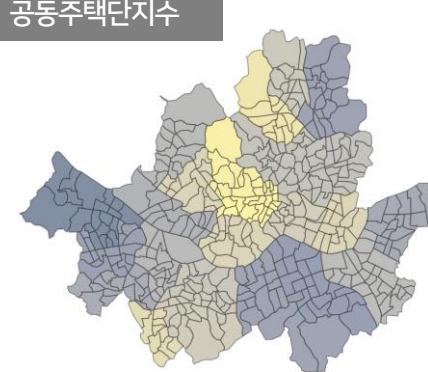
사업체수



민원 점수



공동주택단지수



자치구별 변수 현황 확인을 통해 해당 변수 활성화가 이루어져야 하는 자치구 선정

30년 이상 노후주택수, 자택 충전소 최단거리 평균은 값이 클수록
그 외 변수들은 값이 작을수록



해당 자치구에 개선이 필요함을 알 수 있음

공간회귀 | 결과

변수별 가중회귀계수 자치구 상위 5개

- 사업체수: 강북구 | 도봉구 | 종로구 | 성북구 | 노원구
- 공동주택단지수: 종로구 | 성북구 | 강북구 | 성동구 | 강남구
- 민원 접수: 광진구 | 중랑구 | 강동구 | 동대문구 | 성동구
- 사회환경: 영등포구 | 관악구 | 동작구 | 마포구 | 금천구
- 여성카드소비비율: 마포구 | 영등포구 | 양천구 | 구로구 | 강서구

변수별 현황 자치구 하위 5개

- 사업체수: 도봉구 | 강북구 | 동작구 | 서대문구 | 용산구
- 공동주택단지수: 중구 | 종로구 | 강북구 | 금천구 | 광진구
- 민원 접수: 마포구 | 영등포구 | 송파구 | 서대문구 | 중랑구
- 사회환경: 금천구 | 중랑구 | 용산구 | 도봉구 | 강동구
- 여성카드소비비율: 금천구 | 중구 | 영등포구 | 동대문구 | 서초구

* 표시한 변수 외 다른 변수는 가중회귀계수와 현황 간의 공통된 자치구 찾을 수 없음

해당 변수가 최종 지수에 미치는 영향이 큰 & 현황 분석 결과 개선이 필요한 자치구 선정
= 동일 변수에 대해 회귀계수 값이 크지만, 현황에서 다른 자치구에 비해 값이 작은 자치구 선정

↓

각 변수별 상위 5개 자치구를 추출한 결과, 표시된 자치구가 해당 변수의 개선이 필요함

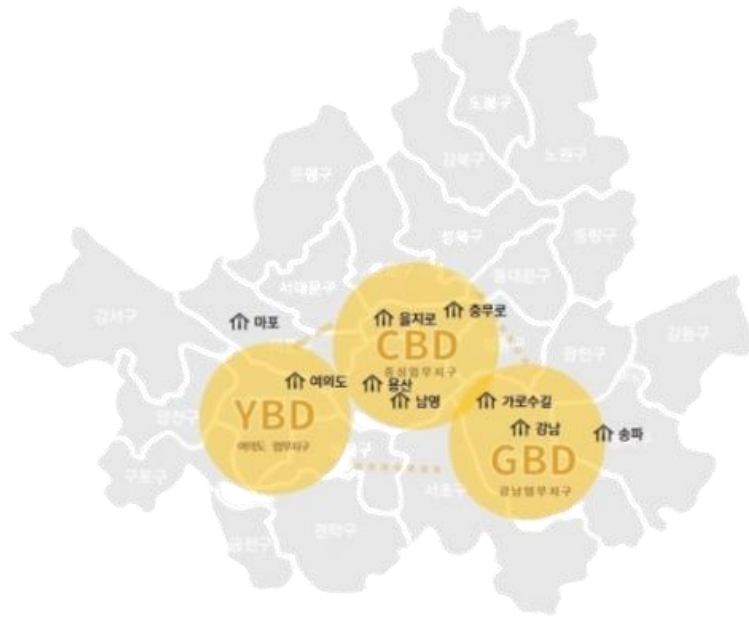


선정된 결과와 현황 자료조사 고려하여
자치구별 활성화 방안 도출

공간회귀 | 인사이트

① 종로구 - 공동주택단지 추가 설립

- 현재 개수가 적으나, 회귀계수의 영향력을 고려하면 추가 설립을 하였을 때 최종 지수에 긍정적인 영향을 미침



종로구는 서울 3대 주요 업무지구로 주거 수요가 높지만
주거 단지는 한정되어 있어 가격이 높음

신아일보 2023.07.06

서울 창신·송인동 신통기획 통해 '2000가구 공동주택' 추진

서울시가 창신·송인동 일대 신속통합기획안을 확정하고
'구릉지형 도심 단지'를 컨셉트로 2000가구 규모 공동주택
조성을 추진한다.
효율적 토지 이용을 위해 대상지 내 저이용·방치 시설을 재
배치하고 공공시설을 고도화한다. ... (후략)

종로구 일대에 공공 주택 설립 계획 추진 과정에 있음

➔ 앞선 분석 결과와 시행 정책 및 2030 서울 생활권 발전 계획을 고려하였을 때, 종로구에 공동주택단지를 설립하였을 때 **최종지수 활성화 기대**

공간회귀 | 인사이트

② 중랑구 - 민원 점수

- 현재 점수가 낮으나, 회귀계수의 영향력을 고려하면 민원 점수가 개선될수록 최종 지수에 긍정적인 영향을 미침

경향신문 2023.10.11

중랑구, 전국 첫 '악취지도' 완성...
하수구 악취 잡는다

신아일보 2023.04.27

중랑구, 고질적 음주 민원 해결 위해
발 벗고 나선다

구에 따르면 주취소란으로 인한 지역주민의 불안감 해소 및
쾌적한 광장 조성을 위해 면목역광장을 금주구역으로
추진중이라고 17일 밝혔다.

2006년 처음 조성된 면목역광장은 중랑구를 대표하는 지하철역
인근 광장이자 복합문화공간임에도 불구하고, 음주와 소음/소란
등 고질적인 문제로 민원이 발생하곤 했다. ...(후략)

중랑구는 민원 발생 해결에 적극적으로 대응하고 있음



국민권익위원회



행정안전부

민원행정 관리 기반 (민원행정 전략 및 체계),
민원행정활동 (민원제도 운영, 국민신문고 민원처리, 고충민원처리),
민원처리성과 (민원만족도)

2022년 행정안전부 시행 '민원서비스 종합평가' 결과에서

중랑구는 가장 낮은 등급인 '마 등급'을 받음

➔ 앞선 분석 결과와 시행 정책 및 2030 서울 생활권 발전 계획을 고려하였을 때, 중랑구 민원 점수를 개선하였을 때 **최종지수 활성화 기대**

인사이트 정리

클러스터링으로 묶인 권역을 통해
권역별 활동지수 향상을 위한 **정책 제언**

| 클러스터 | 인구지 수 평균 | 경제 · 전력지수 평균 | 생활지수 평균 | 문화지수 평균 | 최종 지수 평균 | 자치구 |
|------|-------------|--------------------|------------|------------|----------------|-----------------|
| 5 | 0.07 | 0.23 | 0.05 | 0.34 | 0.70 | 종로구, 중구, 용산구 |

도심권: 「도심주거 확충, 노후주거지 개선, 교통 접근성 강화」 제언

클러스터링 결과와 2030 서울 생활권 발전 계획을 고려하였을 때
도심권에 노후주거지 개선, 교통 접근성 강화 등을 실시한다면
활동지수 활성화 기대 가능

공간회귀 결과 회귀계수와 변수별 현황을 고려해
해당 변수가 낮은 자치구에 **정책 제언**

변수별 가중회귀계수 자치구 상위 5개

- 공동주택단지수: 종로구 | 성북구 | **강북구** | 성동구 | 강남구

변수별 현황 자치구 하위 5개

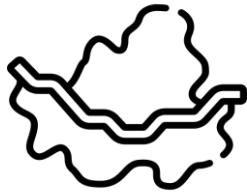
- 공동주택단지수: 중구 | **종로구** | 강북구 | 금천구 | 광진구

종로구: 「공동주택단지 추가 설립」 제언

공간회귀 결과와 시행 정책 등을 고려하였을 때
종로구에 공동주택단지를 추가로 설립한다면
활동지수 활성화 기대 가능

의의 및 유의할 점

활동지수 산출 의의



- ① 쇠퇴된 지역을 판단하기 위한 객관적인 지표
지방 재정 투자 심사 및 정책 수립 시 활용 가능
- ② 서울시 도시 경쟁력 강화 및 지역균형발전 정책 실현

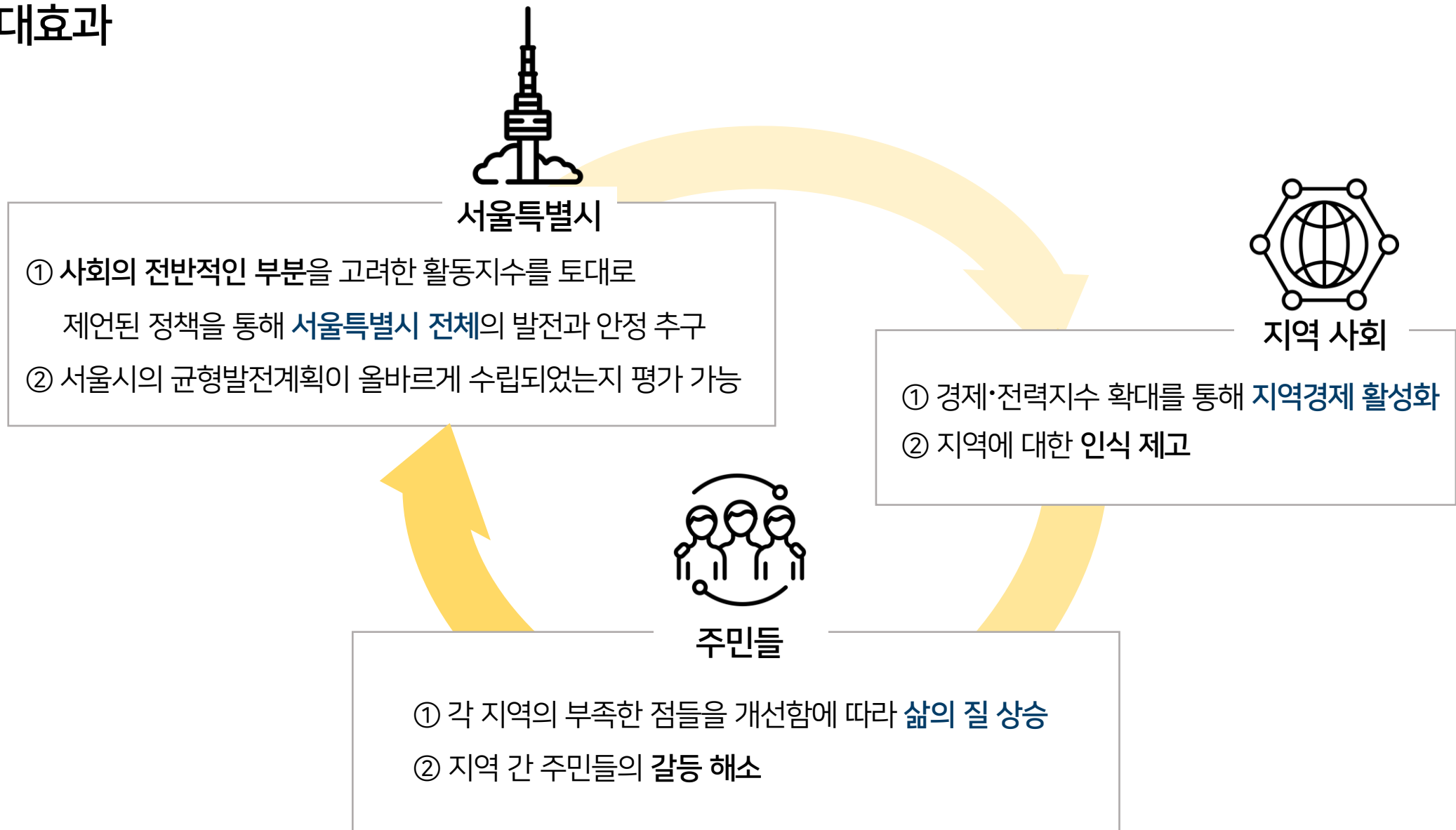
유의할 점

지역 특성을 발전시키고
지역 간 경쟁을 불러 일으킴으로써
지역 발전의 원동력이 되는 지역불균형전략



지나친 전략은 자원수요의 과밀 또는 과소 현상을
불러 일으켜 자원의 효율적 이용 및 국민 통합을 저해

기대효과



기대효과

활동지수의 독창성

- ① 기존의 쇠퇴지표의 문제점을 보완,
다양한 측면을 활용하여 지수 산출
- ② 서울시 내의 발전지수를 살펴보면서,
서울의 지역적 특성을 고려한 차별화된 지수 개발

그동안 집중되어 연구된 수도권-비수도권 격차나
수도권 남, 북부 지역 간 격차가 아닌 서울에 초점!

발전 가능성

- ① 데이터 안심구역 데이터를 활용한 통계적 분석으로
신뢰성 있는 결과 도출
- ② 서울시 뿐만 아니라 다른 도시,
국가 전체의 지역 균형 발전을 위한 지수로 확장 가능
- ③ 지수를 통해 서울시 내 경제적 기회와
도시서비스에 대한 보편적 접근성 제공 가능

감사합니다
