

서울시 도시가스 이용 가구수, 인구, 소득 데이터를 활용한 에너지 소비 특성 분석 및 클러스터링

다차원적 데이터 기반의 서울시 에너지 소비 패턴 심층 분석

서론

서울시는 대한민국의 수도이자 대표적인 대도시로, 거주 인구와 가구의 에너지 소비 패턴이 매우 다양합니다. 특히, 도시가스는 서울 시민의 일상생활에 필수적인 에너지원으로 자리 잡고 있으며, 각 가정의 도시가스 이용 형태는 인구 구조와 소득 분포 등 다양한 요인에 의해 영향을 받습니다. 본 문서에서는 서울시의 도시가스 이용 가구수, 인구, 그리고 소득 데이터를 통합적으로 분석하여 에너지 소비 특성을 파악하고, 이를 기반으로 클러스터링 기법을 활용해 대표적 소비 유형을 도출하고자 합니다.

본 프로젝트는 서울시의 도시가스 에너지소비 특성을 심층적으로 이해하고, 지역별 맞춤 형 에너지 정책 수립에 기여하고자 합니다. 2019년부터 2023년 까지 서울시 자치구별 도시가스 이용가구수, 총 상주인구수, 월 평균 소득 데이터를 통합 분석하여 인구 및 소득 변화가 도시가스 이용에 미치는 영향을 파악하고, 서울시 25개 자치구를 유사한 에너지 소비 특성별로 클러스터링 했습니다.

데이터 구성 및 수집

서울시 에너지 소비 특성 분석을 위해 다음 세 가지 주요 데이터를 활용하였습니다.

- 서울시 도시가스현황, 서울시 상주인구, 서울시 소득소비 통계 등 데이터는 서울 열린데이터광장에서 활용 하였습니다.

에너지 소비 특성 분석

1. 도시가스 이용률과 월평균 소득, 인구의 상관관계



데이터 수집

서울 열린데이터광장에서 도시가스현황, 상주인구, 소득소비 통계 데이터를 수집하여 분석에 활용하였습니다.



분석 방법

수집된 데이터를 통합하여 도시가스 이용률과 월평균 소득, 인구 간의 상관관계를 분석하고 클러스터링 기법을 적용했습니다.

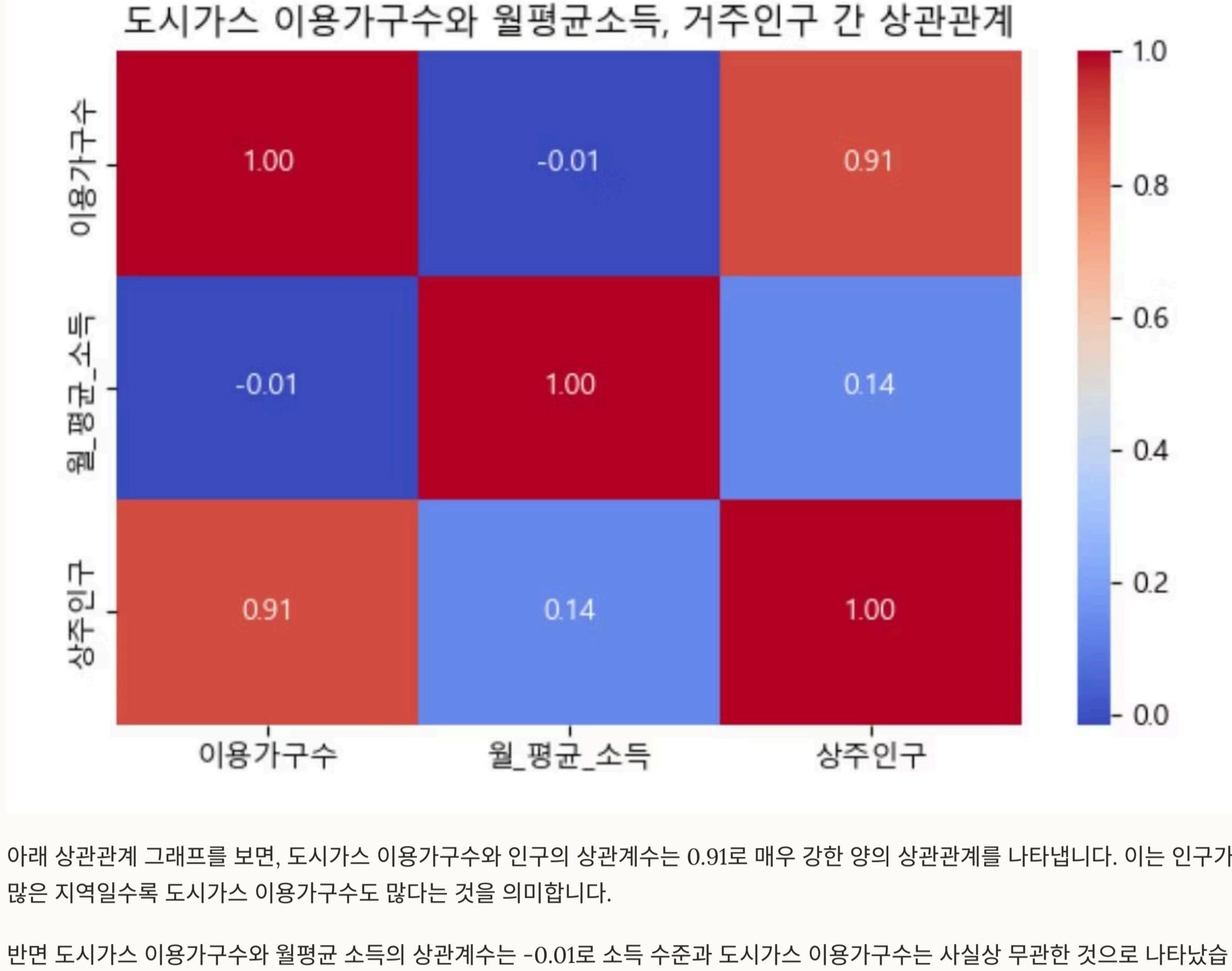


분석 기간

2019년부터 2023년까지 5년간의 데이터를 활용하여 시계열적 변화와 패턴을 파악하였습니다.

도시가스 이용과 인구, 소득의 상관관계

도시가스 이용은 인구 수에 크게 영향을 받는 반면, 소득 수준과는 큰 연관이 없는 것으로 나타났습니다.

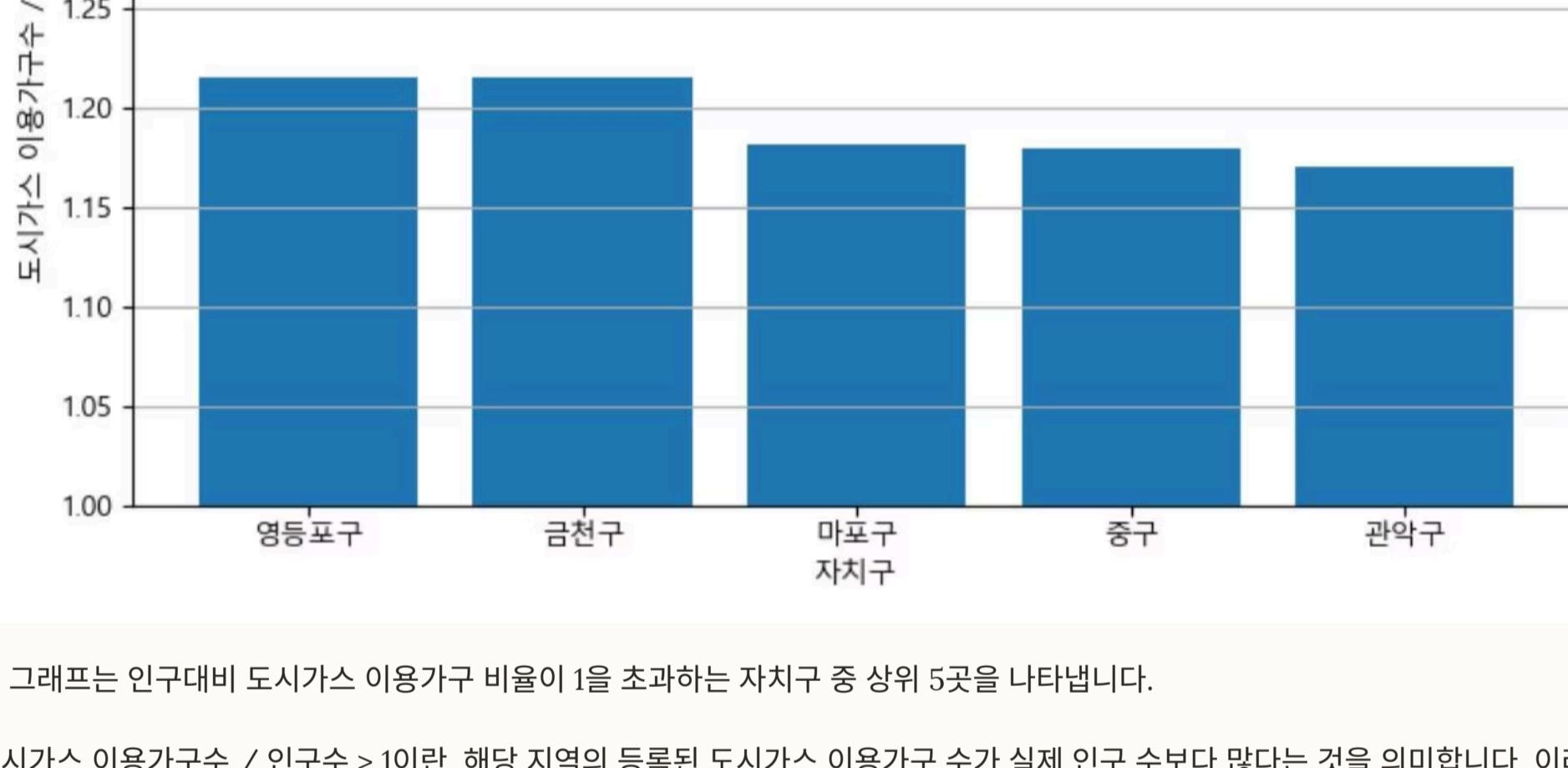


아래 상관관계 그래프를 보면, 도시가스 이용가구수와 인구의 상관계수는 0.91로 매우 강한 양의 상관관계를 나타냅니다. 이는 인구가 많은 지역일수록 도시가스 이용가구수도 많다는 것을 의미합니다.

반면 도시가스 이용가구수와 월평균 소득의 상관계수는 -0.01로 소득 수준과 도시가스 이용가구수는 사실상 무관한 것으로 나타났습니다.

또한 도시가스 이용률과 인구의 상관계수는 -0.63으로 인구가 많을수록 도시가스 이용률은 낮아지는 경향이 있습니다. 이는 인구는 많지만 상대적으로 도시가스 이용가구 비중이 낮은 지역이 존재함을 의미합니다. 이러한 현상은 주로 1인 가구 비중이 높은 지역이거나, 상업, 업무 지구 중심의 자치구 일 수 있습니다.

※ (도시가스 이용률 = 이용가구수 / 거주인구)

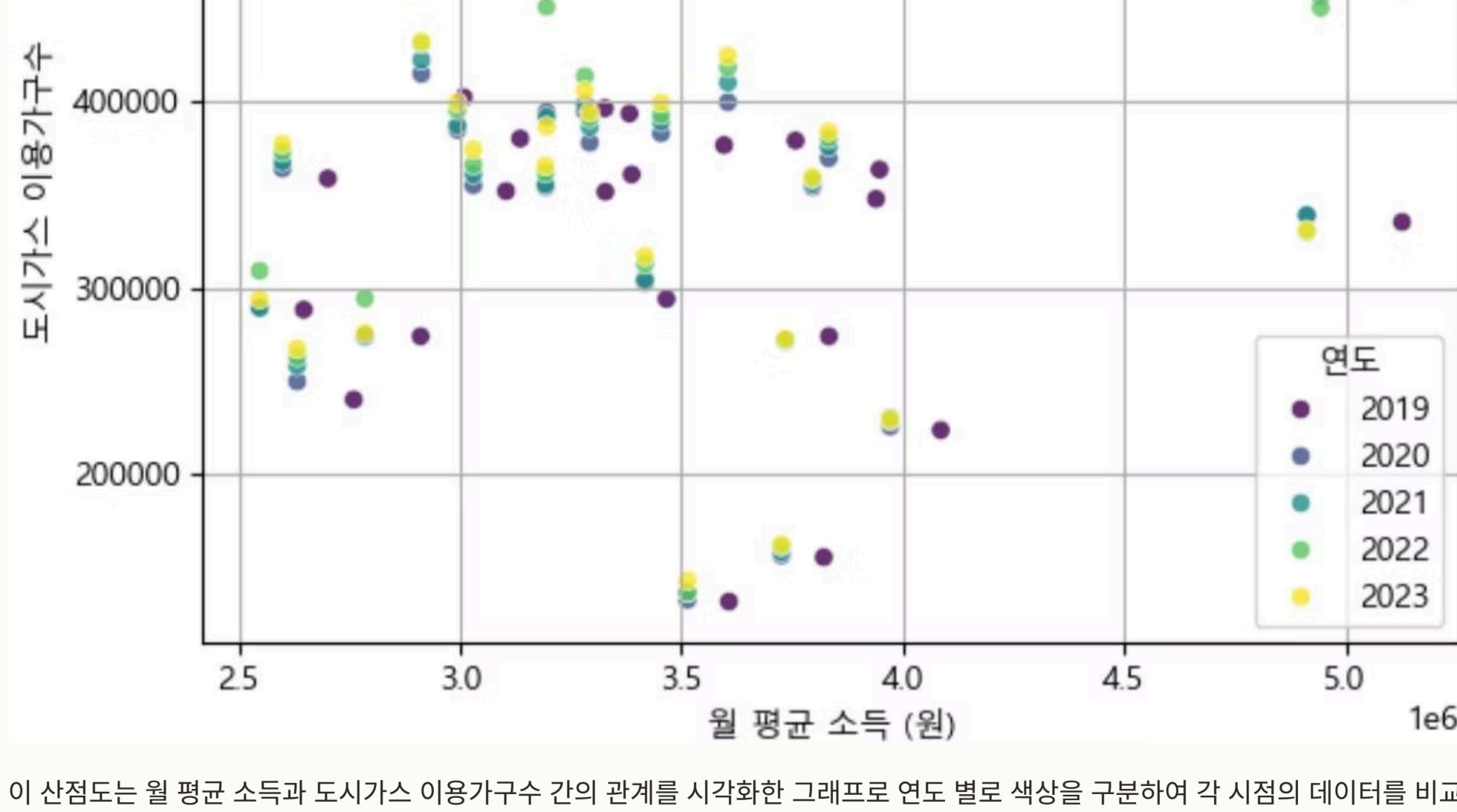


위 그래프는 인구대비 도시가스 이용가구 비율이 1을 초과하는 자치구 중 상위 5곳을 나타냅니다.

도시가스 이용가구수 / 인구수 > 1이란, 해당 지역의 등록된 도시가스 이용가구 수가 실제 인구 수보다 많다는 것을 의미합니다. 이러한 현상은 1인가구 밀집, 비상주 가구 또는 주택 수 통계 기준의 차이 등에 의해 발생 할 수 있습니다.

종합적으로 볼때, 도시가스 수요는 인구 규모에 의해 주로 결정되며, 소득 수준은 도시가스 소비에 유의미한 영향이 미치지 않는다는 결론을 도출할 수 있습니다.

2. 소득 수준과 도시가스 이용가구수 간 관계



이 산점도는 월 평균 소득과 도시가스 이용가구수 간의 관계를 시각화한 그래프로 연도 별로 색상을 구분하여 각 시점의 데이터를 비교하고 있습니다.

전체적으로 볼때, 소득이 증가한다고 하여 도시가스 이용가구수가 뚜렷하게 증가하는 경향은 보이지 않으며, 다양한 소득 구간에서 도시가스 이용가구수가 고르게 분포되어 있음을 알 수 있습니다.

즉 소득 수준과 도시가스 이용가구수 간에 명확한 상관관계는 없는 것으로 판단되며, 이는 상관계수 -0.01이라는 통계적 수치와도 일치합니다.

특히 고소득 지역에서도 도시가스 이용가구수가 상대적으로 적은 경우가 다수 존재하는 점은 이를 더욱 뒷받침합니다.

클러스터링 분석

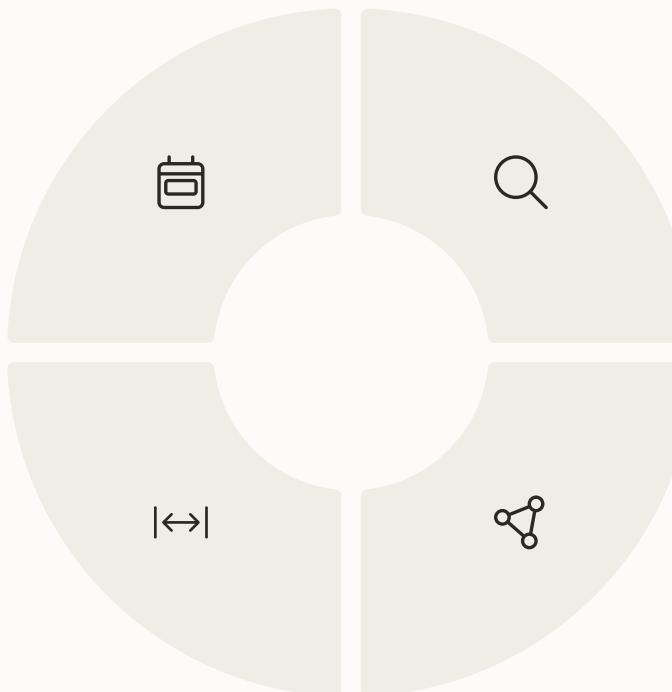
에너지 소비 특성 데이터(도시가스 이용 가구수, 인구, 평균 소득)를 표준화하여 K-means 클러스터링 기법을 적용하였습니다.

데이터 표준화

도시가스 이용 가구수, 인구, 평균 소득
데이터를 표준화하여 분석에 활용

군집별 특성 분석

도출된 군집별 에너지 소비 특성을 분석
하여 정책적 시사점 도출



K-means 알고리즘

유사한 특성을 가진 자치구들을 그룹화
하기 위해 K-means 클러스터링 기법
적용

최적 군집 수 결정

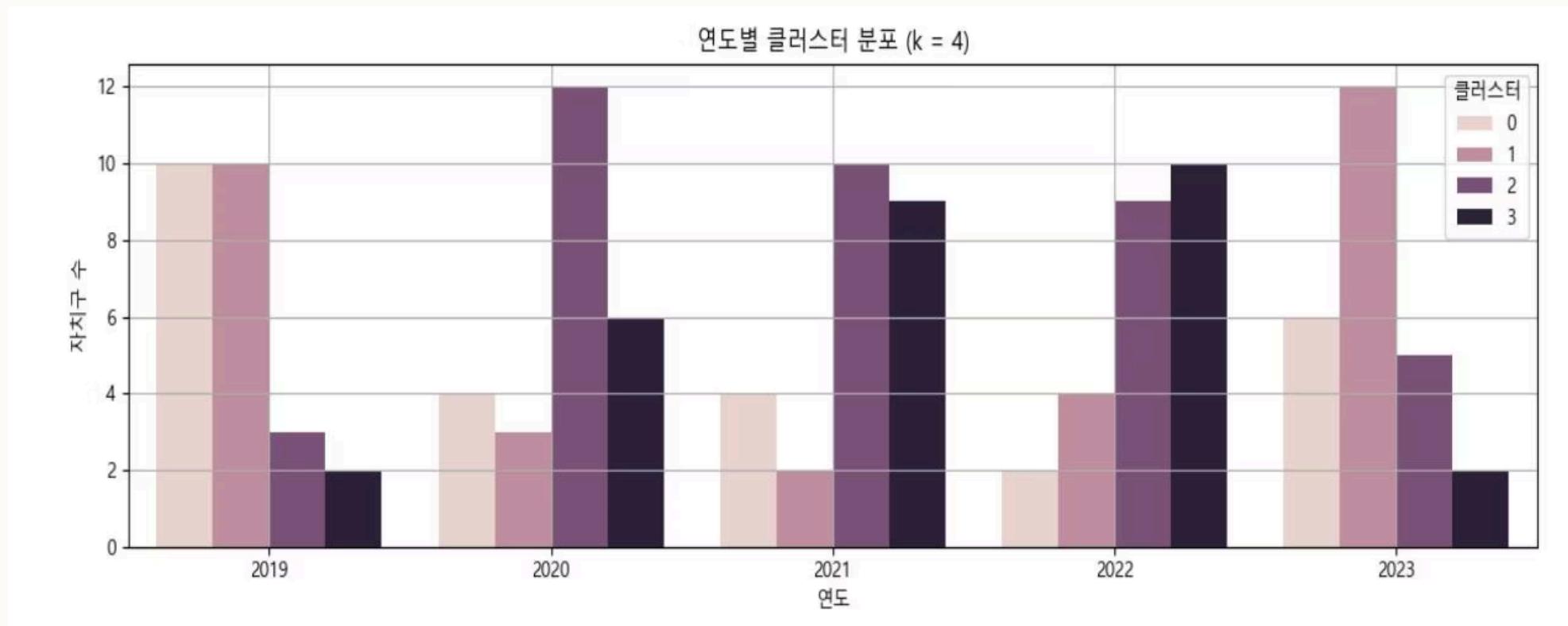
실루엣 계수를 기준으로 2~9개 범위에
서 최적 군집 수 탐색

군집화 방법론

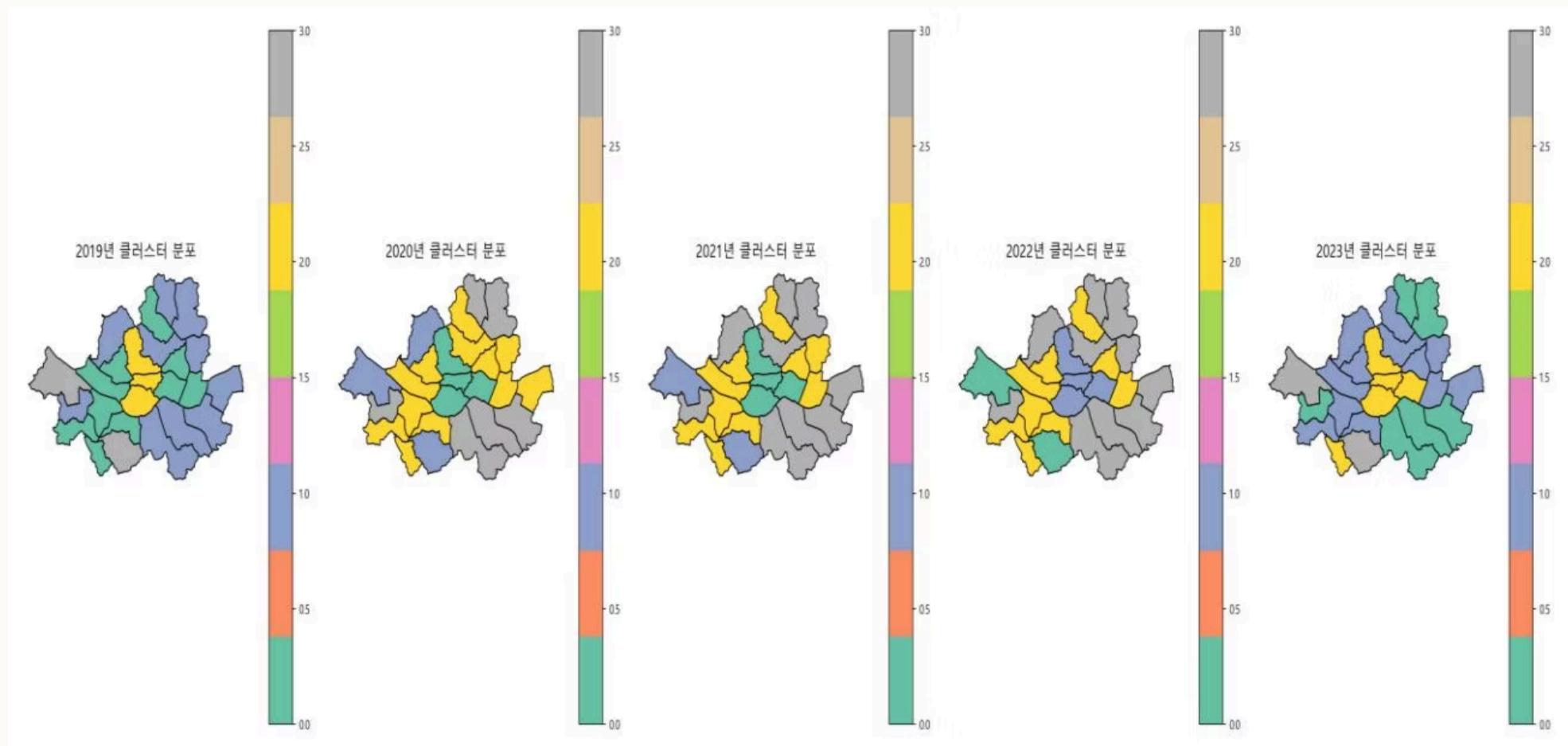
- 군집수(k)는 실루엣 계수를 기준으로 2~9개 범위에서 탐색 하였습니다.

그 결과, 일부 연도에서는 $K = 2$ 에서 가장 높은 실루엣 계수를 보였으나, 이는 단순히 고소비 vs 저소비의 양극단 비교로 실제 도시가스 소비 특성을 충분히 반영하기엔 한계가 있었습니다.

반면, 다수의 연도에서 $k = 4$ 일 때 가장 높은 실루엣 계수를 나타냈고, 도시가스 소비 구조상 자치구를 4개의 유형으로 나누는 것이 가장 의미 있는 군집 분류라고 판단하였습니다. 이에 따라 본 분석에서는 $k = 4$ 를 최적 군집 수로 결정하여 클러스터링을 수행하였습니다.



클러스터별 특성 도출



클러스터	평균 이용가구수	평균 인구대비 이용률	평균 소득대비 이용률	특성
0번	약 34만 가구	0.96	0.10	평균적인 소비수준, 인구대비 안정적
1번	약 30만 가구	0.92	0.08	상대적으로 저소비 클러스터. 소득대비 낮음
2번	약 45만 가구	0.93	0.14	가장 높은 소비 클러스터. 고소득 대비 소비도 높음
3번	약 37만 가구	0.96	0.11	중간 ~ 고소비 인구대비 효율적인 소비 구조

클러스터 2번은 도시가스 소비 수준이 가장 높은 집단으로, 정책적 모니터링이 필요하거나 에너지 절약 유도 대상이 될 수 있습니다.

클러스터 1번은 소득 대비 도시가스 소비가 낮은 집단으로, 에너지 복지지원의 우선순위 대상으로 고려될 수 있습니다.

클러스터 0번과 3번은 평균적이거나 약간 높은 수준의 안정된 소비 패턴을 보이며 비교적 균형 잡힌 도시가스 사용 구조를 가진 집단으로 해석할 수 있습니다.

34만

30만

45만

37만

0번 클러스터 평균 이용
가구수

평균적인 소비수준, 인구대비
안정적인 이용률 0.96

1번 클러스터 평균 이용
가구수

상대적으로 저소비 클러스터,
소득대비 이용률 0.08로 가장
낮음

2번 클러스터 평균 이용
가구수

가장 높은 소비 클러스터, 소득
대비 이용률 0.14로 가장 높음

3번 클러스터 평균 이용
가구수

중간~고소비, 인구대비 효율적
인 소비 구조 (이용률 0.96)

결론 및 정책 제안

1 도시가스 소비와 인구 및 소득의 관계

- 도시가스 이용가구수와 인구간의 분석 결과, 상관계수 0.91로 매우 강한 양의 상관관계가 나타났습니다. 이는 인구수가 도시가스 수요를 결정하는 핵심 요인임을 의미합니다.
- 반면 도시가스 이용가구수와 월평균 소득 간의 상관계수는 -0.01로 사실상 상관관계가 없는 것으로 나타났습니다. 이는 소득 수준이 도시가스 소비에 미치는 영향이 매우 미미하다는 점을 알 수 있습니다.
- 또한 도시가스 이용률과 인구간의 상관계수는 -0.63으로 인구가 많을수록 도시가스 이용률은 낮아지는 경향이 확인되었습니다. 이는 1인가구 비중이 높은 지역이거나 상업, 업무 중심 자치구의 특성이 반영된 결과로 해석할 수 있습니다.

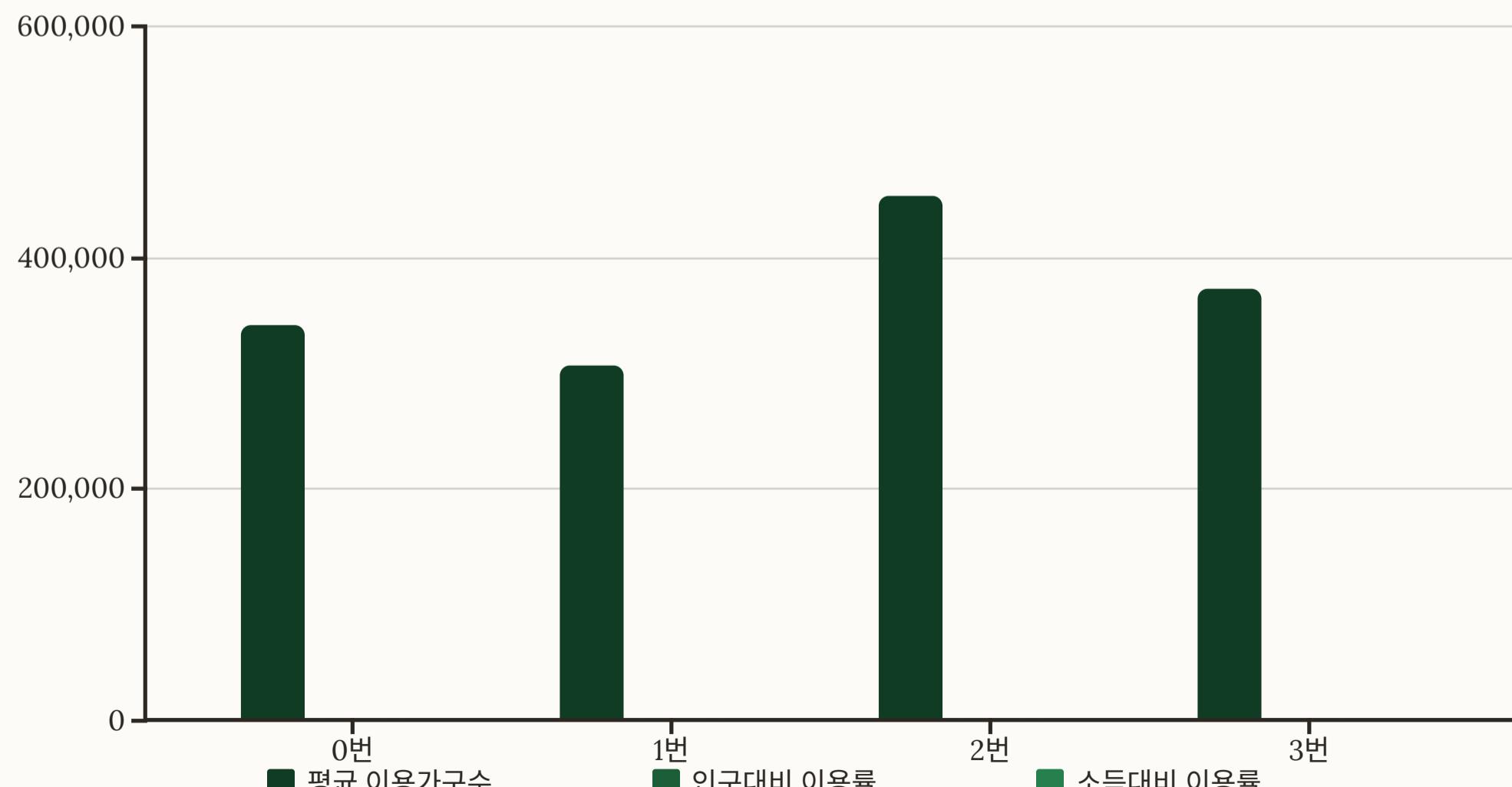
도시가스 소비는 소득 수준보다는 인구 규모와 주거 형태의 영향을 더 크게 받는다는 점 분석을 통해 확인 되었습니다.

따라서 향후 에너지 정책을 설계할 때는 단순한 소득 기준보다는 인구 구조와 가구 형태에 중점을 두고 접근하는 것이 보다 효과적일 것으로 판단됩니다.

2. 클러스터링 결과

실루엣 계수 분석 결과 군집 수를 4개로 설정하는 것이 가장 의미 있는 분류 기준으로 판단되었습니다. 이는 이용가구수, 인구대비 이용률, 소득대비 이용률을 기준으로 서울시 자치구를 4개의 그룹으로 구분하는 방식이 가장 자연스럽고 균형 잡힌 군집 구조를 형성한다는 것을 보여줍니다

클러스터	평균 이용가구수	인구대비 이용률	소득대비 이용률
0번	341,424 가구	0.96	0.10
1번	307,016 가구	0.92 (최저)	0.08(최저)
2번	453,314 가구	0.93	0.14(최고)
3번	373,215 가구	0.96	0.11



클러스터별 정책적 시사점

0번, 3번 클러스터

표준적 관리 모델, 안정적인 수요관리 유지 가능

1번 클러스터

에너지 복지 지원 우선 대상, 공급 사각지대 점검

2번 클러스터

에너지 절약 유도, 고효율 설비 보급, 실시간 모니터링 필요



클러스터별 특성에 맞는 차별화된 에너지 정책 수립이 필요합니다. 특히 1번 클러스터는 에너지 복지 지원이, 2번 클러스터는 에너지 효율화 정책이 중점적으로 필요합니다.



클러스터 특성 분석

각 클러스터의 도시가스 소비 패턴과 특성을 심층 분석



맞춤형 정책 설계

클러스터별 특성에 맞는 차별화된 에너지 정책 수립



정책 실행 및 모니터링

설계된 정책의 효과적 실행과 지속적인 성과 모니터링



피드백 및 정책 개선

정책 효과 분석을 통한 지속적인 개선 및 보완

종합 결론

1 도시가스 소비는 인구 구조와 주거 특성에 큰 영향을 받으며, 단순히 소득 수준으로는 어렵다는 점이 확인되었습니다.

2 클러스터링 기법을 통해 자치구를 4개의 소비 특성그룹으로 분류함으로써 서울시의 에너지 소비 구조를 보다 명확하게 파악할 수 있었습니다.

3 이를 바탕으로, 정책설계 및 지원 방향을 클러스터별로 차별화할 수 있는 근거를 마련하였습니다.

- 예 : 고소비 집단에는 에너지 효율화 유도 정책,
- 저소비 집단에는 에너지 복지 지원 강화 등 차등적 접근 가능

주요 발견점

- 도시가스 이용가구수와 인구의 상관계수: 0.91 (강한 양의 상관관계)
- 도시가스 이용가구수와 월평균 소득의 상관계수: -0.01 (무상관)
- 도시가스 이용률과 인구의 상관계수: -0.63 (음의 상관관계)
- 최적 클러스터 수: 4개 (실루엣 계수 기준)

정책적 함의

향후 에너지 정책 설계 시 단순 소득 기준보다는 인구 구조와 가구 형태에 중점을 두고 접근하는 것이 효과적입니다. 클러스터별 특성에 맞는 차별화된 정책 접근이 필요하며, 특히 고소비 집단(2번)과 저소비 집단(1번)에 대한 맞춤형 정책이 중요합니다.

