

## 서울시 자치구별 주택가격과 인프라 데이터를 통한 동네 특성 분석

김지훈<sup>○</sup>, 백재순<sup>\*</sup>, 김성진(교신저자)<sup>\*</sup>

<sup>○</sup>명지전문대학 ICT융합공학과,

<sup>\*</sup>명지전문대학 ICT융합공학과

e-mail: 2023821028@mjc.ac.kr<sup>○</sup>, {hisoon99, ict214548}@mjc.ac.kr<sup>\*</sup>

## Analysis of Neighborhood Characteristics through Housing Prices and Infrastructure Data for Each Autonomous District in Seoul

Ji-Hoon Kim<sup>○</sup>, Jai-Soon Baek<sup>\*</sup>, Sung-Jin Kim(Corresponding Author)<sup>\*</sup>

<sup>○</sup>Dept. of ICT Convergence Engineering, Myongji College,

<sup>\*</sup>Dept. of ICT Convergence Engineering, Myongji College

### ● 요약 ●

본 논문에서는 자치구별 집 가격과 인프라 데이터를 통한 분석을 기반으로, 저렴한 주택 지역에 입주하는 사람들의 우려와 관련하여 좋은 동네와 안좋은 동네의 차이를 다각도로 조망하고자 한다. DataSet은 서울 열린 데이터 광장과 보건의료 빅데이터 개방 시스템에서 수집한 공공데이터를 활용한다. dependent variable로는 자치구별 인프라 데이터셋을 사용하였으며, independent variable는 자치구별 집 가격을 기반으로 데이터 분석을 수행한다. 본 논문에서는 다양한 분석 기법을 활용하여 모델의 정확도와 신뢰성을 향상시키고, 이를 토대로 동네의 특징과 주거 환경의 차이를 명확히 도출하여 결론을 이끌어내고자 한다.

**키워드:** 집값(Property Prices), 주거환경(Residential Environment), 주변 특성(Neighborhood Characteristics)

### I. Introduction

최근 언론에서는 주택 가격이 낮은 동네가 아이들에게는 거주하기에 부적절한 환경일 수 있다는 주장이 나오고 있다[1]. 이에 대한 논의가 활발히 이루어지고 있는데, 특히 고가의 아파트에 사는 사람들이 주택 가격이 낮은 지역에 살고 있는 이웃들을 차별하는 사례에 대한 보도도 종종 나오고 있다[2]. 이러한 뉴스는 사람들 간의 주거지 선택이 사회적인 차별을 야기할 수 있다는 문제에 대한 관심을 불러일으키고 있다.

이에 더해 좋은 동네와 안좋은 동네의 차이를 파악하기 전에 좋은 동네의 정의를 찾는 것이 중요하다는 점이 강조되었다. 주택의 가치를 결정하는 주요 요소 중 하나로 인프라가 중요하게 작용한다는 기사 내용을 바탕으로[3], 주택 가격이 높은 동네가 더 나은 인프라를 갖추고 있는지를 확인하고자 하는 가설이 제시하고자 한다.

주택 가격이 높은 자치구에서는 어떤 인프라가 더 발전되어 있는지, 특히 안전 측면에서 어떤 차이가 있는지를 알아보기 위해 서울 열린 데이터 광장과 보건의료 빅데이터 개방 시스템에서 수집한 공공데이터를 기반으로 데이터를 수집하고, 이를 토대로 동네의 특징과 주거 환경의 차이를 명확히 도출하여 사회적인 차별 문제에 대한 심층적인

이해를 제공하고자 한다.

### II. Preliminaries

강신혜(2023)에 연구에서는 신도시와 재개발이 가져다줄 새로운 인프라에 기대한 집값 상승에 대해 분석적으로 접근하고 있다[4].

김진우(2023)는 주거환경 인프라가 주택전세가격에 어떤 영향을 미치는지에 대한 연구를 진행하였다. 교통인프라와 녹지환경 인프라를 중심으로 아파트, 오피스텔, 연립다세대에 대한 주택유형별 주택전세가격의 차이를 확인하고 있다. 연구에 따르면 주택전세가격은 지하철과 가까울수록 높아지고, 공원과 가까울수록 낮아지는 경향이 나타나고 있다[5].

조용성(2020)은 서울시 중계동을 대상으로 교육인프라 접근성이 공동주택 가격에 미치는 영향에 대해 살펴보고, 연 결과 학교와의 거리가 공동주택 거래 가격 격차에 가장 큰 영향을 주는 것으로 나타났다[6].

김명연(2019)은 서울시를 대상으로 근린환경이 공동주택가격에 미치는 영향을 살펴보고, 단지 내 세대수, 카페 및 베이커리 수, 사설학원가가 공동주택가격 상승의 요인으로 밝혀졌다[7].

이러한 선행 연구들을 바탕으로 본 연구는 서울시를 대상으로 자치구별로 집값에 영향을 주는 다양한 요인들을 분석해보는 연구를 진행하였다.

### III. The Proposed Scheme

#### 1. 분석 기법

##### 1.1 집 가격과 인프라 변수 간의 상관관계 분석

분석 기법 중 하나로 집 가격과 인프라 변수 간의 상관관계를 살펴보고, 평당가를 기준으로 한 집 가격과 역 개수 간에는 양의 상관관계가 나타났다. 이는 대중교통 시설이 더 많은 지역에서 평당가가 가격이 상승하는 경향을 보여주며, 이는 해당 지역의 교통 편의성이 주택가격에 긍정적인 영향을 미친다는 것을 시사한다.

또한, 편의점 수와도 양의 상관관계가 확인되었다.

이는 주변 편의시설이 늘어날수록 주거 환경의 편의성이 높아지며, 이로써 주택가격이 상승하는 경향을 보여준다. 따라서 이러한 결과들은 지역의 대중교통 및 편의시설 확보가 주택시장에 미치는 긍정적인 영향을 데이터 기반으로 확인할 수 있다는 의미를 지니고 있다.

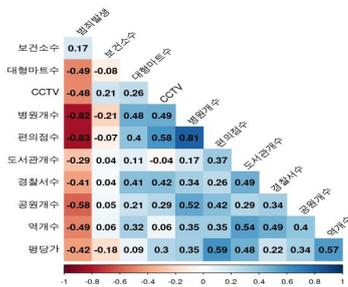


Fig. 1. System Architecture

##### 1.2 회귀분석의 등분산성 및 정규성 검정

등분산성 및 정규성을 검정한 결과, 회귀분석에서는 통계적으로 유의한 문제가 나타나지 않았다. Q-Q Plot과 ncvTest를 활용한 등분산성 검정에서는 등분산성 가정이 만족되었으며, 셰이퍼로-윌크 정규성 검정에서는 종속 변수의 분포가 정규분포를 따르지 않을 가능성이 낮게 나타났다. 총론적으로, 이 모델은 안정적이며 신뢰성 있는 예측을 제공할 것으로 판단된다.

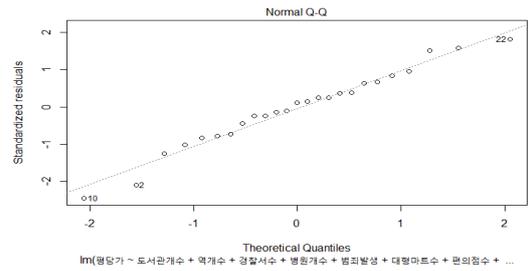


Fig. 2. Homoscedasticity and normality tests in regression analysis

##### 1.3 시각화

회귀분석을 기반으로 집값에 직접적으로 영향을 미치는 변수를 파악하기 위해 변수들을 순위별로 시각화하였다.

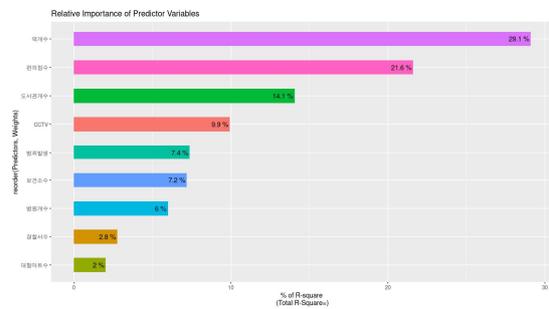


Fig. 3. Visualization

##### 1.4 로지스틱 검정

집값에 따른 변수들의 영향을 더 자세히 살펴보기 위해 로지스틱 회귀분석에서 호스머-렘쇼 분석을 실시했다. 이 분석에서 얻은 p-value는 0.9464로, 이는 모형의 적합도가 통계적으로 유의미하게 높다는 것을 의미한다. 따라서 집값이 높음 또는 낮음에 따른 변수들의 영향을 효과적으로 파악할 수 있는 우수한 모델임을 확인하였다.

##### 1.5 요인 분석 전

요인분석 수행 전 데이터의 적합성을 평가하기 위해 KMO(Kaiser-Meyer-Olkin) 검정과 Bartlett 검정을 실시하였다. KMO 검정 결과, 측정된 변수들 간에 0.59의 공통성이 존재함을 나타내어 데이터가 요인분석에 적합한지 확인하였다. 더불어 Bartlett 검정에서는 변수들 간에 선형 상관성이 확인되어, 변수들 간의 분산 차이가 없어 요인분석을 신뢰성 있게 진행할 수 있다는 결론을 도출하였다. 이러한 검정 결과를 통해 요인분석이 신뢰성 있게 이루어질 수 있음을 확인하고 분석을 진행하였다.

##### 1.6 요인분석

KMO 및 Bartlett 검정을 통해 요인분석을 신뢰성 있게 진행할 수 있다는 결론을 도출한 후, 요인 분석을 시행하였다. 요인 분석의 과정은 주성분 분석을 통한 시각화 및 요인 개수 설정 후, 고유값과 Scree plot을 활용하여 요인 개수를 결정하고, Loadings 값 중 0.5

이상인 것들을 추출하였다. Fig. 4에서는 이러한 결과를 시각화한 요인지표를 제시하였다. 그림에서 빨간 표시된 부분은 적재량을 나타내는데, 이를 통해 어떤 변수들이 어떤 요인에 속하며 개별 관측치의 요인 점수를 확인할 수 있었다.

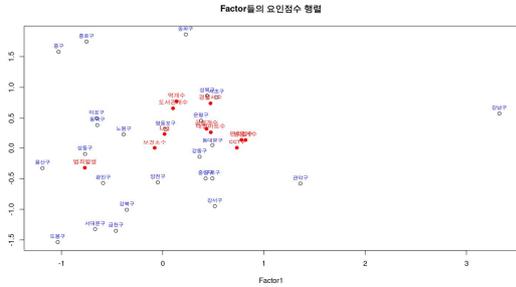


Fig. 4. System Architecture

### 1.7 요인분석 및 비교분석

집값에 따른 요인 및 변수의 차이를 비교하기 위해 요인 분석과 비교 분석을 수행했다. 각 요인과 변수를 기준으로 평균 이하는 0, 평균 이상은 1로 구분하여 분포를 시각화하였다. Fig. 5에서는 병원 수, 범죄 발생 수, 편의점 수를, Fig. 6에서는 경찰서 수, 역 개수, 도서관 수를, 그리고 Fig. 7에서는 CCTV 수를 시각화한 결과이다. 이를 통해 집값이 높은 지역과 낮은 지역 간의 다양한 인프라 및 시설의 차이를 명확하게 확인할 수 있다.

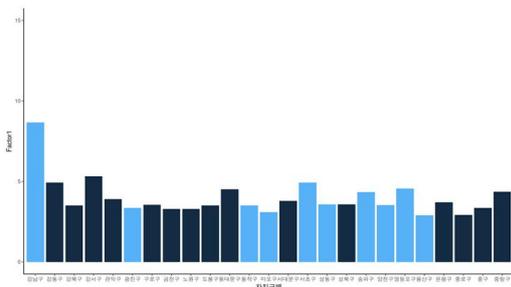


Fig. 5. Factor analysis and comparative analysis -1

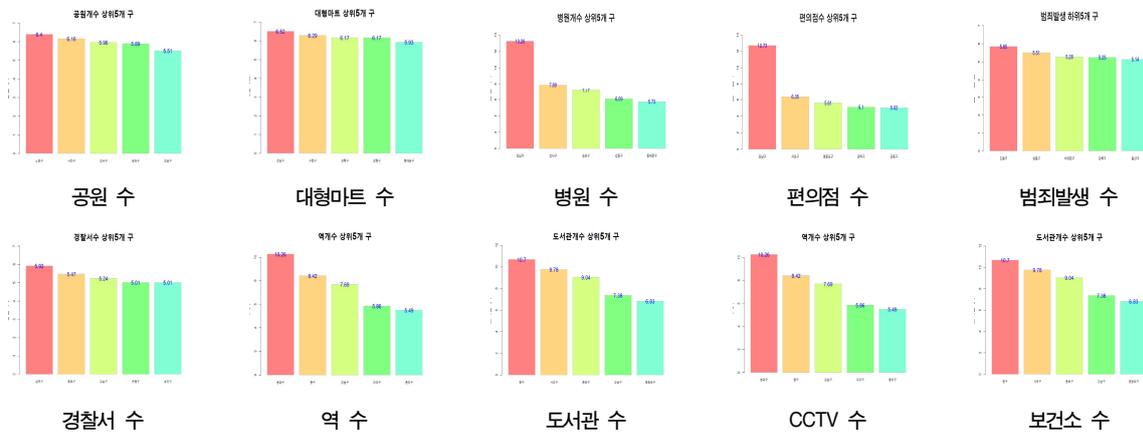


Fig. 9. Visualize the top 5 wards

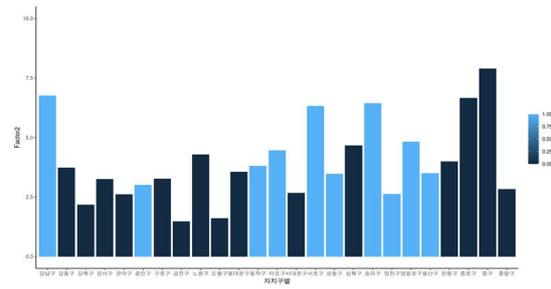


Fig. 6. Factor analysis and comparative analysis -2

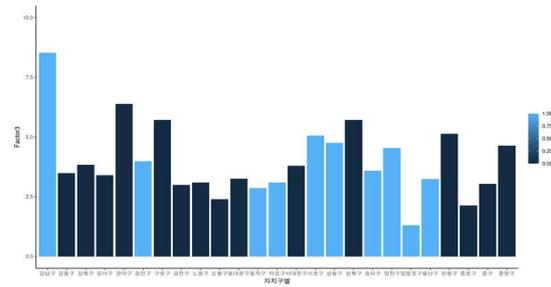


Fig. 7. Factor analysis and comparative analysis -3

### 1.8 요인분석 및 비교분석 결과 시각화

공원, 대형마트, 병원, 편의점, 범죄발생, 경찰서, 역, 도서관, CCTV, 보건소의 수에 따른 각 자치구의 집값을 시각화해보았다[Fig.9].

공원개수는 노원구, 서초구, 강서구, 송파구, 강남구는 공원 개수가 높게 나타났으며 특히 강남구는 높은 집값에도 불구하고 공원 확보에 신경을 쓰고 있는 것으로 보인다.

대형마트는 강남구, 구로구, 성북구, 강동구, 동대문구는 대형마트의 입점이 활발하게 이루어진 것으로 나타났다.

병원개수는 강남구, 강서구, 송파구, 강동구, 동대문구는 병원의 개수가 높게 나타났으며, 높은 집값 지역에서는 의료 인프라에 대한 투자가 크게 이루어져 있다.

편의점개수는 강남구, 서초구, 영등포구, 관악구, 강동구는 편의점의 개수가 높게 나타났다. 상대적으로 더 높은 소득 수준의 지역에서는

24시간 편리한 상점의 입점이 활발한 것으로 보인다.

범죄발생수는 도봉구, 성동구, 서대문구, 강북구, 용산구는 범죄 발생 수가 낮은 지역에 속한다.

경찰서수는 성북구, 종로구, 강남구, 은평구, 송파구는 경찰서 수가 높게 나타났습니다. 안전을 강조하는 지역에서는 경찰서의 배치가 활발한 것으로 보인다.

역의 개수는 송파구, 중구, 강남구, 마포구, 종로구는 지하철 역 개수가 많은 편에 속한다. 대중교통 인프라의 편리성이 높은 동네로 해석된다.

도서관의 개수는 중구, 서초구, 종로구, 강남구, 영등포구는 도서관 개수가 많은 편에 속해, 교육 환경이나 문화 인프라에 대한 투자가 큰 것으로 나타난다.

CCTV의 개수는 강남구, 관악구, 구로구, 성북구, 은평구는 CCTV의 개수가 높은 편에 속해 있습니다. 안전을 강조하고 범죄 예방에 투자가 이루어지고 있는 것으로 보인다.

보건소의 개수는 중구, 노원구, 성동구, 성북구, 은평구는 보건소의 개수가 많은 편에 속해 있다. 주민들의 건강을 챙기는 인프라에 투자가 이루어지고 있는 것으로 해석된다.

apartment prices - targeting the academy-dense area of Eunhaeng intersection in Junggye-dong. Korean Journal of Urban Design, Urban Design, 21(6), 5-14.

- [7] Myeong-Yeon Kim, & Eun-Jeong Kim. (2019). The impact of neighborhood environment on apartment prices: Targeting Seoul Metropolitan City. Journal of the Korean Society of Urban Geography, 22(1), 163-175.

#### IV. Conclusions

종합적으로 판단하면, 집값이 높다고 해서 반드시 지역의 우수성이 보장되는 것은 아니며, 반대로 비교적 저렴한 동네에서도 높은 생활 품질을 제공하는 지역이 있음을 명확히 확인하였다. 이는 단순한 집값이 지역의 품질을 결정하는 데에 있어서는 제한적인 시각임을 시사한다. 다양한 분석 결과를 통해 나타난 이러한 흥미로운 현상은 집값과 지역의 품질 간에 단순한 원인과 결과의 관계가 아니라, 복잡한 상호작용과 영향을 받는다는 점을 강조하고 있다.

#### REFERENCES

- [1] <https://www.weeklyseoul.net>
- [2] <https://www.segye.com>
- [3] <https://www.mk.co.kr/news/realestate/9329493>
- [4] Shinye Kang. (2023). Protection of housing associations where urban (re)development frequently occurs: receiving education and appreciation Behavioral strategies of residents in the original downtown of Seongnam City and the paradox of value increase. Oral History Research, 14(2), 183-235.
- [5] Jinwoo Kim. (2023). The impact of residential environment infrastructure on housing rental prices. Housing Research, 31(4), 5-31.
- [6] Cho Yong-seong, Lee Jae-hyun, & Kim Hyeong-bo. (2020). The impact of accessibility to education infrastructure on