**[빅데이터 분석 결과 보고서]**

**MZ 세대를 위한 1인 가구**

**생활 패턴 분석 및 추천 서비스**

**김용헌**

**목 차**

[1. 분석 배경 및 의도 4](#_Toc143707605)

[2. 분석 목표 4](#_Toc143707608)

[3. 분석 데이터 5](#_Toc143707610)

[3.1. 데이터 수집 5](#_Toc143707611)

[3.1.1. 인구 통계 및 유입 정보 5](#_Toc143707612)

[3.1.2. 부동산 가격 및 거래 정보 6](#_Toc143707613)

[3.1.3. 생활 편의시설 정보 6](#_Toc143707614)

[3.2. 데이터 전처리 7](#_Toc143707615)

[3.2.1. 서울시 부동산 데이터 전처리 7](#_Toc143707616)

[3.2.2. 서울시 편의시설 데이터 전처리 8](#_Toc143707617)

[3.2.3. 전처리 데이터 통합 8](#_Toc143707618)

[3.3. 데이터 정합성 검정 9](#_Toc143707619)

[4. 분석 과정 및 결과 9](#_Toc143707621)

[4.1. 데이터 선택 기준 9](#_Toc143707622)

[4.2. 분석 방법 10](#_Toc143707623)

[4.3. 분석 도구 및 언어 10](#_Toc143707624)

[4.4. 변수 선택 10](#_Toc143707625)

[4.5. 분석 결과 및 시각화 10](#_Toc143707626)

[4.5.1. 서울시 아파트 부동산 분석 결과 10](#_Toc143707627)

[4.5.2. 서울시 아파트 부동산 데이터 시각화 14](#_Toc143707628)

[5. 추천 알고리즘 및 결과 30](#_Toc143707629)

[5.1. 추천 알고리즘 개요 30](#_Toc143707630)

[5.2. 주요 과정 31](#_Toc143707631)

[5.3. 추천 알고리즘 결과 31](#_Toc143707632)

[5.3.1. 샘플 데이터 셋 31](#_Toc143707633)

[5.3.2. 추천 결과 32](#_Toc143707634)

[6. 활용방안 33](#_Toc143707635)

[7. 참고문헌 34](#_Toc143707640)

1. **분석 배경 및 의도**
   1. **분석 배경**

서울시 내 청년층, 특히 MZ 세대 중 상당수가 수도권으로의 이동과 함께 1인 가구 생활을 시작하고 있다. 학업, 취업 등의 필요로 서울과 경기 지역으로 집중적으로 이동하는 이들은 새로운 주거 환경에 적응하는 과정에서 다양한 정보에 대한 필요성을 느끼고 있다. 그러나, 주거 형태의 선정에 있어 필요한 정보가 부족하거나, 다양한 요인을 고려하여 선택하기가 어려운 상황이다. 본 보고서는 이러한 청년층의 필요성에 부응하여 서울시의 주거 관련 정보를 분석하고, 최적의 주거 형태를 추천하기 위한 빅데이터 분석을 시행한다.

* 1. **분석 의도**

1인 가구 증가는 현대 사회의 주요 사회적 변화 중 하나로 대두되고 있다. 이 변화는 주로 비자발적인 요인, 예를 들어 학업이나 취업으로 인한 독립 생활에 의해 주도되고 있다. 반면, 경제적 요인 등 외부적 상황에 따라 자발적인 1인 가구 시작 의향은 지속적인 하락세를 보이고 있다. 그럼에도 MZ 세대의 1인 가구 생활에 대한 의향은 높아지는 추세이며, 특히 20대의 경우 가장 높은 의향을 보이고 있다.

이러한 흐름을 고려하면, 주거 형태에 따른 부동산 가격, 편의시설, 인구 밀집도 등의 데이터를 분석하여, 청년층의 주거 선택에 대한 정보 제공 서비스는 이들의 주거 문화에 큰 영향을 미칠 것으로 예상된다. 따라서, 본 프로젝트는 사용자의 경제 상황, 선호도 등을 반영하여 서울시 내에서 최적의 주거 형태를 추천하는 것을 목표로 한다.

1. **분석 목표**
   1. **서울시 내 주거형태 별 부동산 가격 분석**

* 서울시 내 주거형태 (아파트, 오피스텔, 다세대주택)에 따른 부동산 가격의 평균 및 분포를 파악한다.
* 주거 형태에 따른 전세, 매매, 월세의 계약 형태별 가격 변동성과 특징을 자치구별로 조사하고 분석한다. 주거 선택 시 계약 형태에 따른 예상 비용 및 변동 특성을 파악한다.
* 각 자치구의 주거 가격을 서울시 전체 평균과 비교하여 분석한다. 상대적인 주거 가격 위치를 파악하고, 해당 자치구의 주거 가격 형성 특성을 이해한다.
* 사용자로부터 입력 받는 주거형태, 계약 형태, 가용한 보증금, 월세 금액, 주거환경에 대한 우선순위 등의 데이터를 기반으로 사용자의 주거 관련 환경 및 선호도를 상세히 파악한다.
* 서울시 내 각 행정동의 현재의 주거형태, 계약 형태별 가격 분포, 주거환경 특성(지하철역, 편의시설, 주요 공원 등)에 관한 데이터베이스를 구축한다.
* 사용자의 선호도와 행정동 데이터베이스의 정보를 비교하여, 사용자의 선호와 가장 잘 매칭되는 행정동을 찾아내는 알고리즘을 설계한다.

1. **분석 데이터**
   1. **데이터 수집**
      1. **인구 통계 및 유입 정보**

* 본 연구에서는 통계지리정보서비스에서 제공하는 인구 총 조사 데이터를 활용하여 청년 세대의 1인 가구 인구 수를 파악한다. 20세부터 35세의 청년 세대 중 1인 가구로 거주 중인 인구수를 중점적으로 수집하였고, 이를 통해 서울시 내에서 청년 세대의 주거 분포 및 주거 선호를 상세히 파악할 수 있다.

또한, 서울시의 생활 이동 데이터를 활용하여 인구의 전입 및 전출 패턴을 분석하였다. 인구의 이동 패턴은 해당 지역의 경제적 활성화, 주거 환경, 교통 편의성 등 다양한 요인에 영향을 받는다. 이 데이터를 통해 어느 지역이 청년 세대에게 인기가 있으며, 어떤 지역이 청년 세대의 주거 및 생활 환경으로 선호되는지에 대한 통찰을 얻을 수 있었다. 특히, 인구 유입량을 통해 특정 지역의 성장세나 인기도, 그리고 청년 세대의 이동 경로와 주거 선호를 더욱 구체적으로 파악할 수 있다.

* 연령 & 주거 형태별 지역별 인구 파악 – [통계청]
* 서울 생활 이동 데이터 – [서울시 열린데이터광장]
  + 1. **부동산 가격 및 거래 정보**
* 부동산 가격 및 거래 정보는 주거 선택에 있어 결정적인 영향을 미치는 요소 중 하나이다. 이를 분석하기 위해 부동산통계정보 및 서울시 부동산 실제 거래가 Data을 활용하였다. 특히, 서울 지역구의 주택 형태(아파트, 오피스텔, 다세대 주택)별 매매, 전세, 월세 가격 데이터를 중점적으로 수집하였다. 이를 통해 서울시 내에서의 부동산 시장 동향, 가격 변동성, 그리고 청년 세대의 주거 선택에 따른 예상 비용을 상세히 파악한다. 또한, 실시간으로 거래되는 부동산 데이터를 통해 최신의 부동산 시장 상황과 트렌드를 실시간으로 파악한다.
* 서울 부동산 가격 정보- [한국부동산원 부동산통계정보시스템]
* 서울시 부동산 실제 거래가 정보 – [서울시 열린데이터광장]
  + 1. **생활 편의시설 정보**
* 주거 선택 시 주요 시설 및 생활 환경은 큰 영향을 미친다. 이를 위해 서울시에서 제공하는 주요 공원 현황, 지하철역 주소, 대규모점포 인허가 정보, 응급실 위치 정보 API을 활용하여 데이터를 수집한다. 이들 데이터를 통해 서울시 내의 주요 시설과 서비스의 분포, 접근성, 그리고 다양성을 평가할 수 있다. 특히, 청년 세대의 주거 선호와 관련하여 교통 편의, 쇼핑 및 엔터테인먼트 시설, 응급의료 서비스, 그리고 문화적 활동을 즐길 수 있는 공간의 위치와 특성을 상세히 파악할 수 있다. 이를 통해 청년 세대가 주거 선택 시 중요하게 고려하는 생활 환경의 특성과 패턴을 분석한다.
* 서울시 주요 공원 현황 – [서울시 열린데이터광장]
* 서울교통공사 지하철역 주소 정보 – [서울시 열린데이터광장]
* 서울시 대규모점포 인허가 정보 – [서울시 열린데이터광장]
* 서울시 응급실 위치 정보 – [서울시 열린데이터광장]
  1. **데이터 전처리**
     1. **서울시 부동산 데이터 전처리**
* 본 프로젝트에서는 주소 정보를 상세하게 처리하여 자치구와 행정동 정보를 명확히 추출했다. 이 과정을 통해 서울시 내 각 지역의 정확한 위치와 거주 현황을 분석할 수 있는 기반을 마련한다.
* 분석의 핵심에 집중하기 위해 불필요한 정보를 제거한다. 자치구 코드, 법정동 코드, 지번 구분 명 등의 열은 분석의 주요 대상이 아니었기에 해당 정보를 제거하여 데이터의 구조를 최적화한다. 또한, 주거 형태별로 데이터를 재구조화하는 작업을 수행한다. 아파트, 오피스텔, 다세대주택 등의 주거 형태에 따라 별도의 데이터베이스 테이블을 구축하였고, 이를 통해 각 주거 형태별 특성과 거래 동향을 보다 명확하게 분석할 수 있도록 한다.
* 모든 matrix의 지리적 범위를 통일하기 위해 행정동 기반의 데이터를 법정동 기준으로 변환한다. 이 과정은 서로 다른 데이터 소스 간의 지리적 일관성을 확보하기 위해 반드시 필요하고, 이를 통해 정확한 지역별 분석이 가능하다.
  + 1. **서울시 편의시설 데이터 전처리**
* 생활 편의시설 데이터 프레임의 전처리에서도 유사하게 분석에 필요한 핵심 정보를 중점적으로 다뤘다. 주소와 시설 명 외의 불필요한 열은 데이터의 명료성과 효율성을 위해 제거한다.
* 데이터 프레임의 주소 정보는 자치구와 행정동 단위로 상세하게 처리한 후 마찬가지로 행정동을 법정동으로 mapping하여 변환한다. 이를 통해 서울시 내 각 지역별로 편의시설의 분포와 특징을 정확하게 분석할 수 있는 기반을 마련한다.
  + 1. **전처리 데이터 통합**
* 데이터 분석에 있어서 여러 데이터 프레임의 정보를 하나로 통합하는 과정이 필요하다. 이를 위해 python의 데이터 분석 라이브러리인 pandas를 활용하였고, Google Colab 환경에서 진행하였다.
* 먼저, 아래와 같이 필요한 라이브러리를 불러온 후, pandas의 merge 함수를 사용하여 '자치구'와 '행정동'을 기준으로 여러 데이터 프레임들을 통합하였다.
* 통합한 데이터 프레임을 통해, 서울시 내 각 지역에 대한 다양한 정보를 하나의 통합된 데이터 프레임을 통해 활용할 수 있도록 하여 효율성을 향상하였다. 이를 통해, 각 지역의 인구 통계, 부동산 정보, 생활 편의시설 정보 등을 함께 고려하여 보다 통합적이고 정밀한 분석을 수행할 수 있게 되었다.
  1. **데이터 정합성 검정**

통합 데이터 프레임을 분석한 결과, 특정 행정동에서 아파트, 오피스텔, 다세대주택의 금액 데이터에 결측치가 존재함을 확인하였다. 데이터의 완전성과 정확성을 위해 적절한 방법을 통해 결측치를 처리하였다.

* + 1. **결측치 파악 및 중앙값 대체**

서울시 내의 행정동에 대해 주거 형태에 따른 전세, 매매, 월세 금액 데이터에 결측치가 존재함을 확인하였다. 서울시의 전체 행정동에서 다양한 항목의 결측치를 포괄적으로 집계하려 했으나, 그 분포와 양을 정확히 파악하는 것이 복잡하다고 판단하여 다른 결측치 처리 방법을 찾게 되었다.

결측치 처리의 다양한 방법 중에서 중앙값으로 대체하는 방식을 선택하게 된 주된 이유는 두 가지다. 첫째, 중앙값은 극단치에 영향을 상대적으로 적게 받아 데이터의 왜곡을 최소화할 수 있다. 둘째, 아파트, 오피스텔, 다세대주택과 같은 주거 유형의 금액 데이터에서는 각 행정동에 따른 거래 금액이 크게 다를 수 있어 중앙값 활용이 분석의 정확성을 높일 수 있다는 점을 고려하였다. 따라서 중앙값을 활용하여 결측치를 대체하였고, 이를 통해 데이터의 정확성과 신뢰성을 향상시켰다.

1. **분석 과정 및 결과**
   1. **데이터 선택 기준**

분석에 사용된 데이터는 최근의 부동산 가격 현황을 반영하기 위해 2020년부터 2023년 6월까지의 데이터를 선택하였다. 최신의 시장 동향을 반영하고, 주요 부동산 정책 변화를 포함하고 있어 중기적 변동성과 경향을 충분히 파악할 수 있다고 판단하여 해당 기간의 데이터를 수집하였다.

* 1. **분석 방법**
* Z-점수 계산

분석의 첫 단계로, 각 자치구의 부동산 가격의 표준화된 z-점수를 계산하였다. 이 방법은 데이터의 원래 값이 평균에서 얼마나 떨어져 있는지를 표준편차 단위로 나타내어, 서로 다른 스케일의 값을 비교할 때 효과적이다.

* 백분위 점수 및 등급 분석

데이터의 상대적 위치를 파악하기 위해 백분위 점수를 도입하였다. 추가로, 사분위수 기반의 등급 분석을 통해 각 자치구의 부동산 가격 분포와 중심 경향성을 직관적으로 이해할 수 있도록 하였다.

* 1. **분석 도구 및 언어**

분석은 Python 언어를 사용하여 수행하였습니다. Python은 데이터 분석 및 시각화에 강력한 라이브러리를 보유하고 있다. 특히, pandas, matplotlib, 및 seaborn 라이브러리를 활용하여 데이터 전처리, 분석, 및 시각화를 진행하였다.

* 1. **변수 선택**

변수 선택 시, 부동산 시장의 가격 동향을 반영하는 핵심 지표들을 중심으로 하였다. 각각의 주거형태의 매매가, 전세, 보증금, 및 월세 등의 변수들이 포함되었다.

* 1. **분석 결과 및 시각화**
     1. **서울시 아파트 부동산 분석 결과**

|  | **자치구** | **아파트 매매가** | **아파트 전세** | **아파트 월세보증금** | **아파트 월세** |
| --- | --- | --- | --- | --- | --- |
| **0** | 강남구 | 2.796429 | 2.228265 | 2.035164 | 2.246283 |
| **1** | 강동구 | -0.059953 | -0.196517 | 0.235707 | -0.728156 |
| **2** | 강북구 | -0.989346 | -1.347926 | -1.266508 | -0.570088 |
| **3** | 강서구 | -0.106686 | -0.676461 | -0.131497 | -0.029126 |
| **4** | 관악구 | -0.645197 | -0.620403 | -0.447755 | -0.945971 |
| **5** | 광진구 | -0.275968 | 0.405277 | -0.231203 | 0.159961 |
| **6** | 구로구 | -0.800463 | -0.863906 | -0.263970 | -0.784427 |
| **7** | 금천구 | -0.976200 | -0.909357 | -0.345802 | -1.182825 |
| **8** | 노원구 | -0.938506 | -1.263783 | -1.334853 | -1.165621 |
| **9** | 도봉구 | -1.119498 | -1.472157 | -1.497212 | -0.457251 |
| **10** | 동대문구 | -0.515339 | -0.518768 | -0.849849 | -0.253562 |
| **11** | 동작구 | 0.187806 | 0.538215 | 0.449271 | -0.222607 |
| **12** | 마포구 | 0.232074 | 0.704870 | 0.429816 | 0.349573 |
| **13** | 서대문구 | -0.385970 | -0.394795 | -0.415435 | 0.413601 |
| **14** | 서초구 | 2.476793 | 2.357690 | 2.661331 | 1.254123 |
| **15** | 성동구 | 0.459463 | 0.541515 | 0.813062 | 1.444218 |
| **16** | 성북구 | -0.590198 | -0.392604 | -0.398789 | -0.684355 |
| **17** | 송파구 | 0.524634 | 0.671552 | 0.899384 | 0.011903 |
| **18** | 양천구 | -0.007800 | 0.010956 | 0.824450 | 0.289290 |
| **19** | 영등포구 | 0.073489 | 0.118198 | -0.147368 | -0.336977 |
| **20** | 용산구 | 1.637325 | 1.296439 | 0.575980 | 2.754803 |
| **21** | 은평구 | -0.726693 | -0.769905 | -0.841309 | -1.026518 |
| **22** | 종로구 | 0.127789 | 0.808794 | 1.112542 | 0.438680 |
| **23** | 중구 | 0.582689 | 0.644429 | -0.891949 | 0.038711 |
| **24** | 중랑구 | -0.960674 | -0.899617 | -0.973209 | -1.013661 |

**[표 1-1] 서울시 아파트 금액 표준점수**

**표 [1-1]의 표준점수 결과를 백분위 점수로 변환하였다. (소수점 한 자리)**

|  | **자치구** | **아파트 매매가** | **아파트 전세** | **아파트 월세보증금** | **아파트 월세** |
| --- | --- | --- | --- | --- | --- |
| **0** | 강남구 | 99.7% | 98.7% | 97.9% | 98.8% |
| **1** | 강동구 | 47.6% | 42.2% | 59.3% | 23.3% |
| **2** | 강북구 | 16.1% | 8.9% | 10.3% | 28.4% |
| **3** | 강서구 | 45.8% | 24.9% | 44.8% | 48.8% |
| **4** | 관악구 | 25.9% | 26.7% | 32.7% | 17.2% |
| **5** | 광진구 | 39.1% | 65.7% | 40.9% | 56.4% |
| **6** | 구로구 | 21.2% | 19.4% | 39.6% | 21.6% |
| **7** | 금천구 | 16.4% | 18.2% | 36.5% | 11.8% |
| **8** | 노원구 | 17.4% | 10.3% | 9.1% | 12.2% |
| **9** | 도봉구 | 13.1% | 7.0% | 6.7% | 32.4% |
| **10** | 동대문구 | 30.3% | 30.2% | 19.8% | 40.0% |
| **11** | 동작구 | 57.4% | 70.5% | 67.3% | 41.2% |
| **12** | 마포구 | 59.2% | 76.0% | 66.6% | 63.7% |
| **13** | 서대문구 | 35.0% | 34.6% | 33.9% | 66.0% |
| **14** | 서초구 | 99.3% | 99.1% | 99.6% | 89.5% |
| **15** | 성동구 | 67.7% | 70.6% | 79.2% | 92.6% |
| **16** | 성북구 | 27.8% | 34.7% | 34.5% | 24.7% |
| **17** | 송파구 | 70.0% | 74.9% | 81.6% | 50.5% |
| **18** | 양천구 | 49.7% | 50.4% | 79.5% | 61.4% |
| **19** | 영등포구 | 52.9% | 54.7% | 44.1% | 36.8% |
| **20** | 용산구 | 94.9% | 90.3% | 71.8% | 99.7% |
| **21** | 은평구 | 23.4% | 22.1% | 20.0% | 15.2% |
| **22** | 종로구 | 55.1% | 79.1% | 86.7% | 67.0% |
| **23** | 중구 | 72.0% | 74.0% | 18.6% | 51.5% |
| **24** | 중랑구 | 16.8% | 18.4% | 16.5% | 15.5% |

**[표 1-2] 서울시 아파트 금액 백분위 환산**

* 백분위수 기준 설정

우선, 각 변수의 20%, 40%, 60%, 80% 백분위수를 계산하여 데이터의 분포를 5개 구간으로 나눈다.

* 등급 부여 기준:

5 등급: 값이 20% 백분위수 이하일 경우

4 등급: 값이 20% 초과, 40% 이하일 경우

3 등급: 값이 40% 초과, 60% 이하일 경우

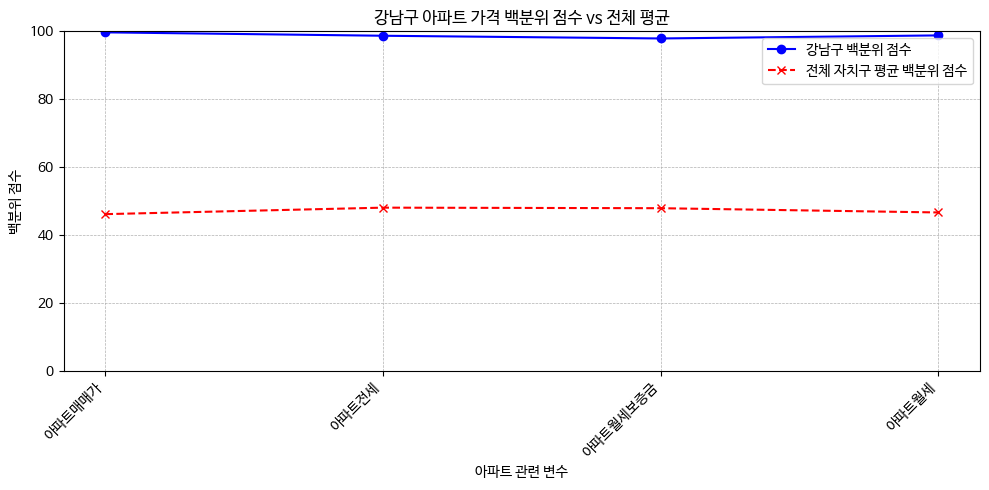
2 등급: 값이 60% 초과, 80% 이하일 경우

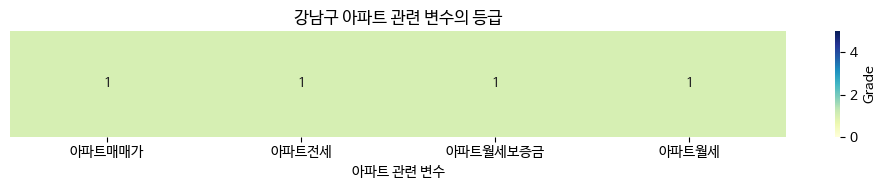
1 등급: 값이 80% 초과일 경우

**[표 1-2]의 백분위 환산 데이터를 위에 언급한 백분위수에 대해 등급을 부여하였다.**

|  | **자치구** | **아파트 매매가** | **아파트 전세** | **아파트 월세보증금** | **아파트 월세** |
| --- | --- | --- | --- | --- | --- |
| **0** | 강남구 | 1 | 1 | 1 | 1 |
| **1** | 강동구 | 3 | 3 | 2 | 4 |
| **2** | 강북구 | 5 | 5 | 5 | 4 |
| **3** | 강서구 | 3 | 4 | 3 | 3 |
| **4** | 관악구 | 4 | 4 | 4 | 5 |
| **5** | 광진구 | 3 | 2 | 3 | 2 |
| **6** | 구로구 | 4 | 4 | 3 | 4 |
| **7** | 금천구 | 5 | 5 | 3 | 5 |
| **8** | 노원구 | 5 | 5 | 5 | 5 |
| **9** | 도봉구 | 5 | 5 | 5 | 4 |
| **10** | 동대문구 | 4 | 4 | 4 | 3 |
| **11** | 동작구 | 2 | 2 | 2 | 3 |
| **12** | 마포구 | 2 | 1 | 2 | 2 |
| **13** | 서대문구 | 3 | 3 | 4 | 2 |
| **14** | 서초구 | 1 | 1 | 1 | 1 |
| **15** | 성동구 | 2 | 2 | 2 | 1 |
| **16** | 성북구 | 4 | 3 | 4 | 4 |
| **17** | 송파구 | 1 | 2 | 1 | 3 |
| **18** | 양천구 | 3 | 3 | 1 | 2 |
| **19** | 영등포구 | 2 | 3 | 3 | 3 |
| **20** | 용산구 | 1 | 1 | 2 | 1 |
| **21** | 은평구 | 4 | 4 | 4 | 5 |
| **22** | 종로구 | 2 | 1 | 1 | 1 |
| **23** | 중구 | 1 | 2 | 5 | 2 |
| **24** | 중랑구 | 5 | 5 | 5 | 5 |

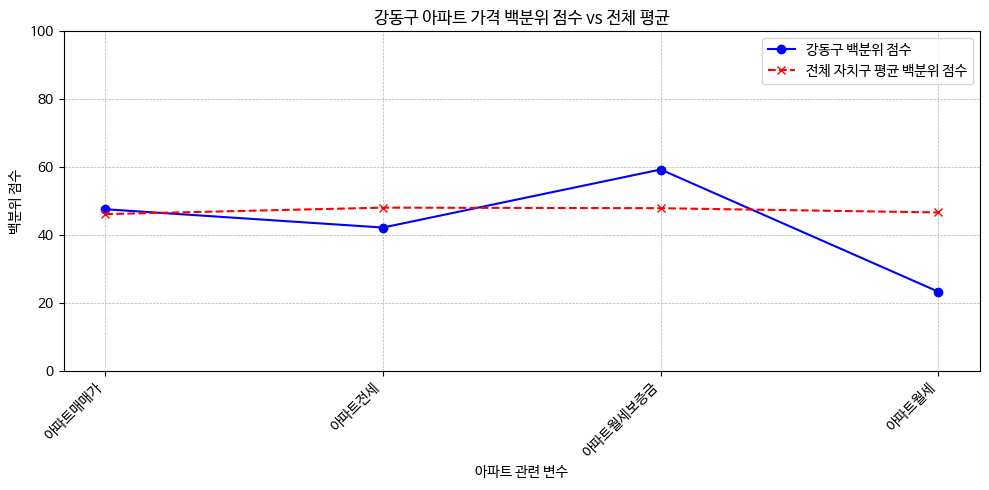
* + 1. **서울시 아파트 부동산 데이터 시각화**
* 강남구

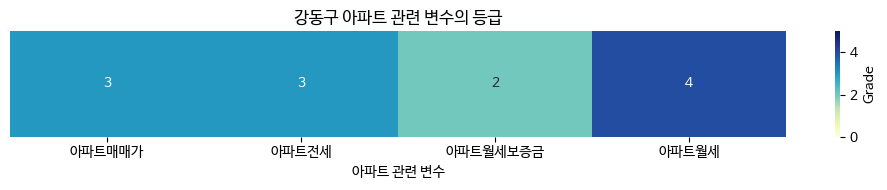




[그래프 1-1] 강남구 아파트 가격 백분위 점수 및 등급

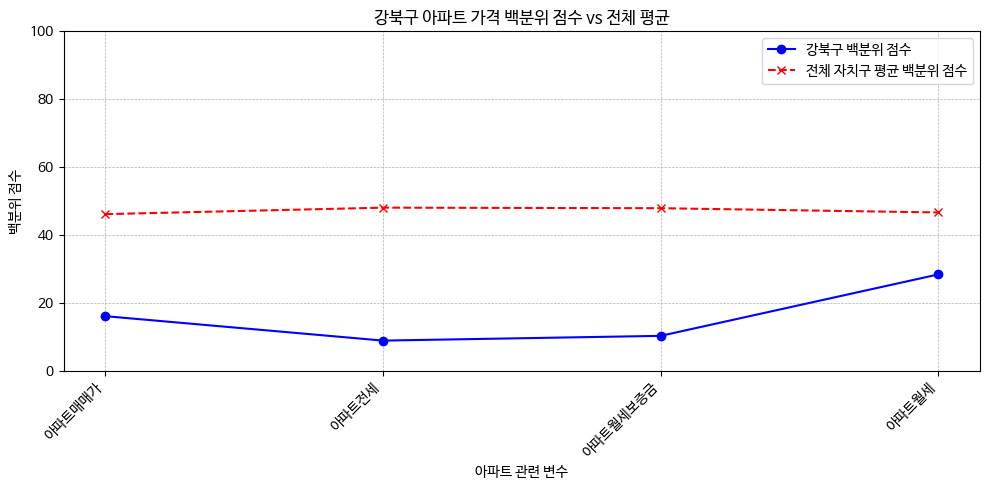
* 강동구

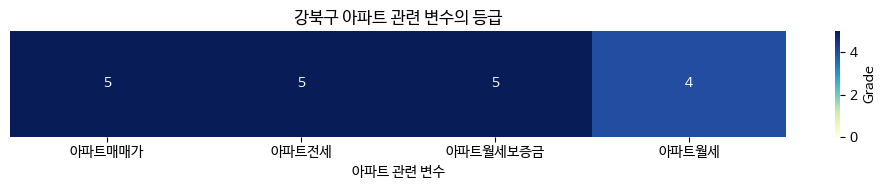




[그래프 1-2] 강동구 아파트 가격 백분위 점수 및 등급

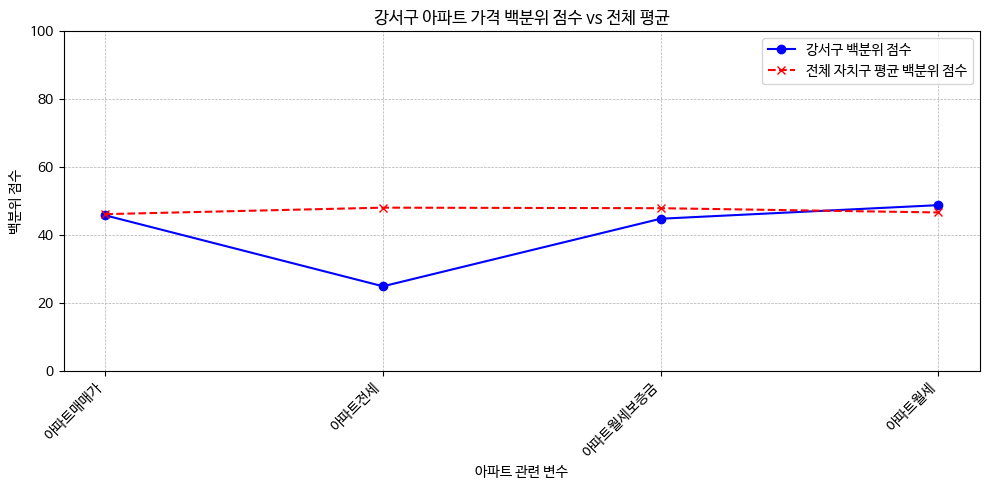
* 강북구

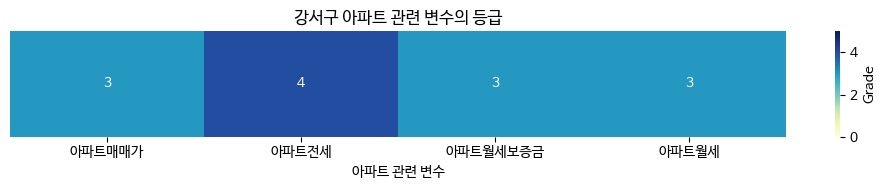




[그래프 1-3] 강북구 아파트 가격 백분위 점수 및 등급

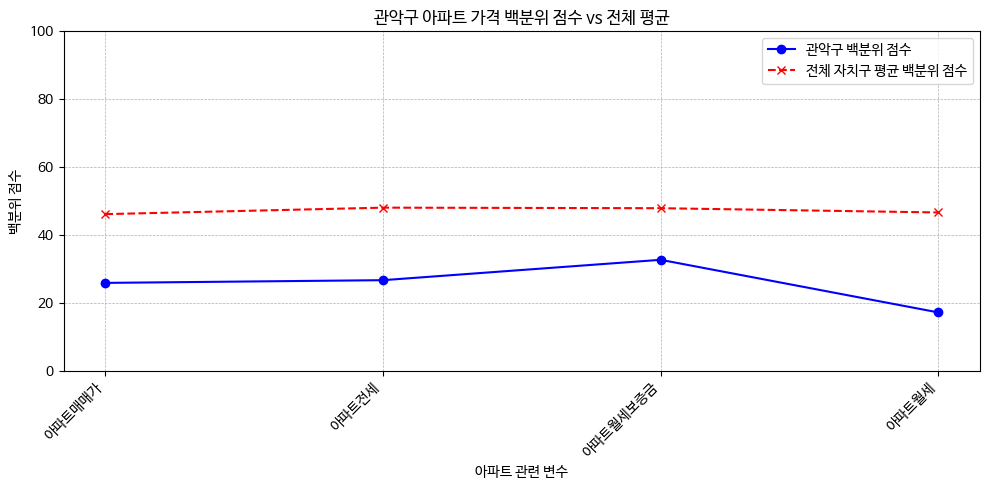
* 강서구

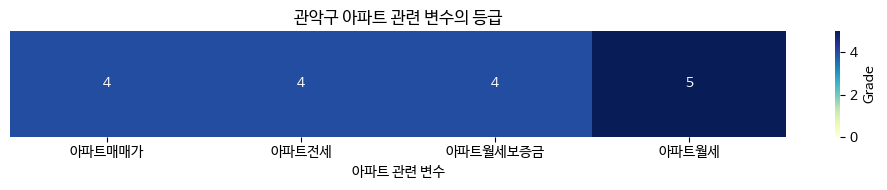




[그래프 1-4] 강서구 아파트 가격 백분위 점수 및 등급

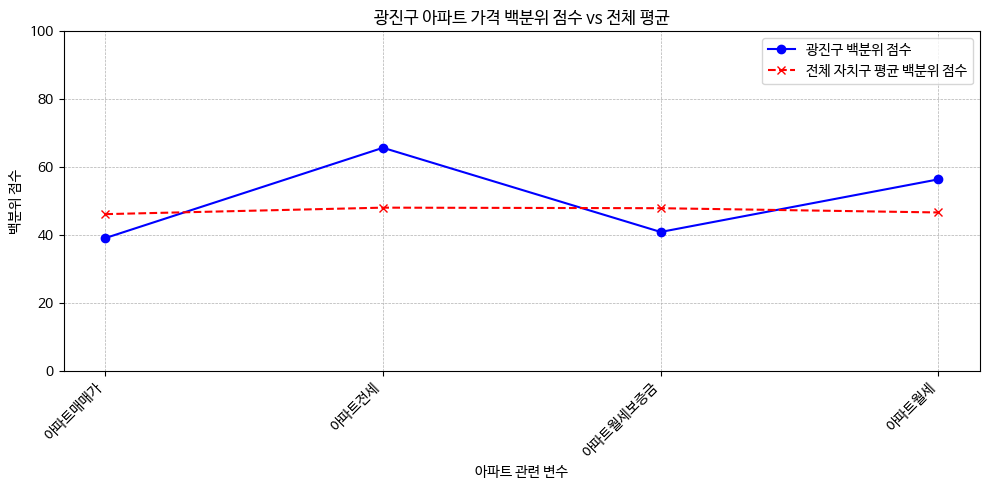
* 관악구

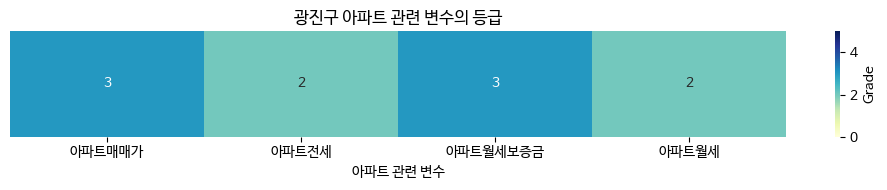




[그래프 1-5] 관악구 아파트 가격 백분위 점수 및 등급

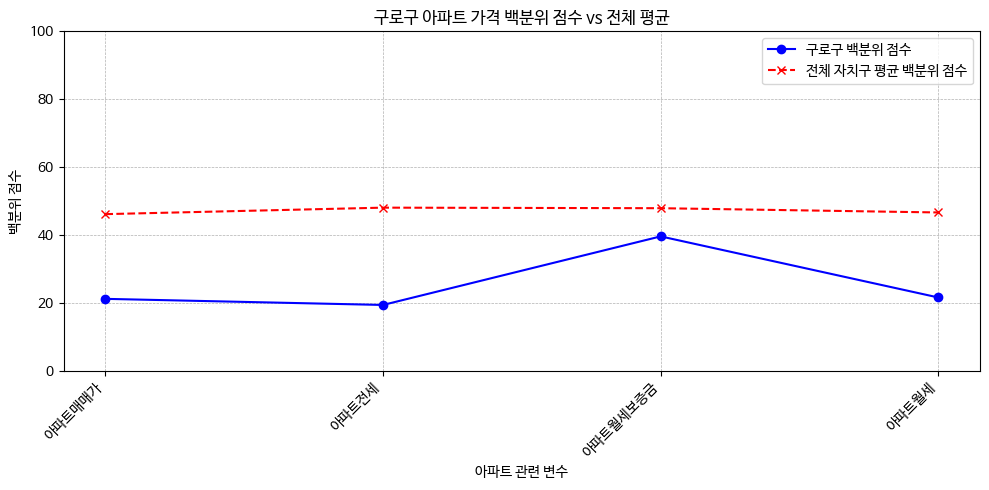
* 광진구

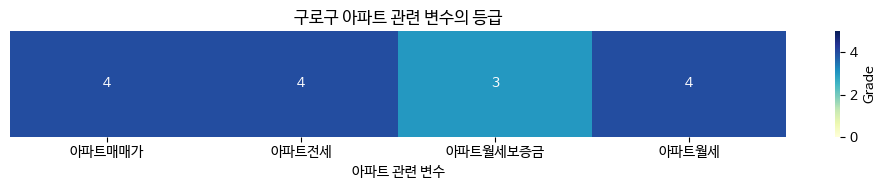




[그래프 1-6] 광진구 아파트 가격 백분위 점수 및 등급

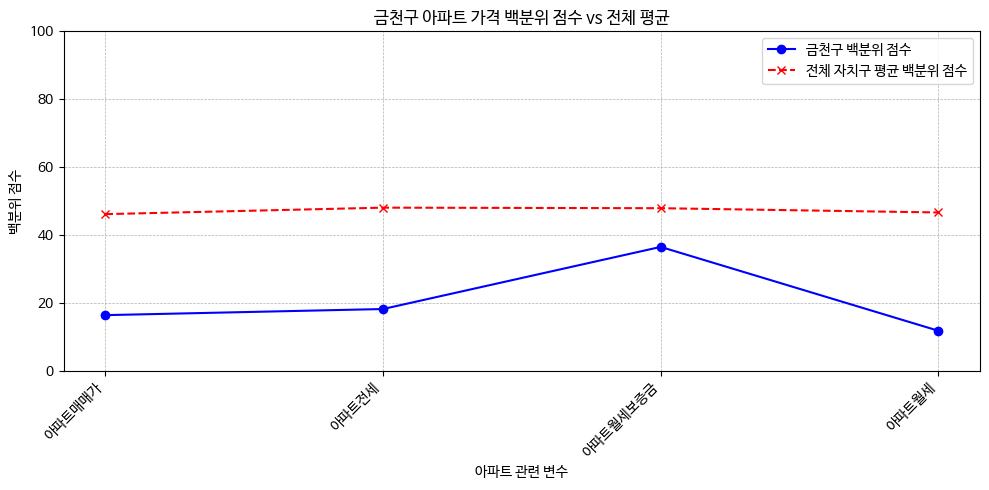
* 구로구

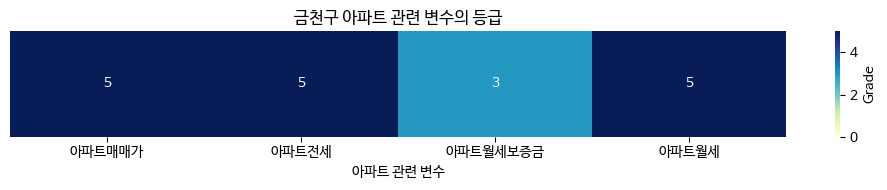




[그래프 1-7] 구로구 아파트 가격 백분위 점수 및 등급

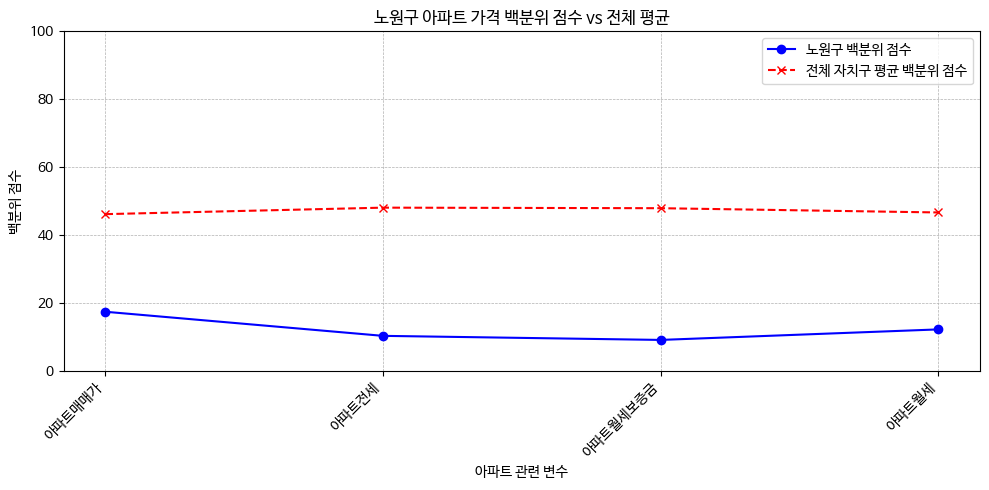
* 금천구

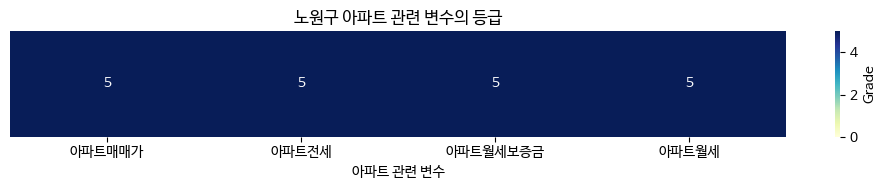




[그래프 1-8] 금천구 아파트 가격 백분위 점수 및 등급

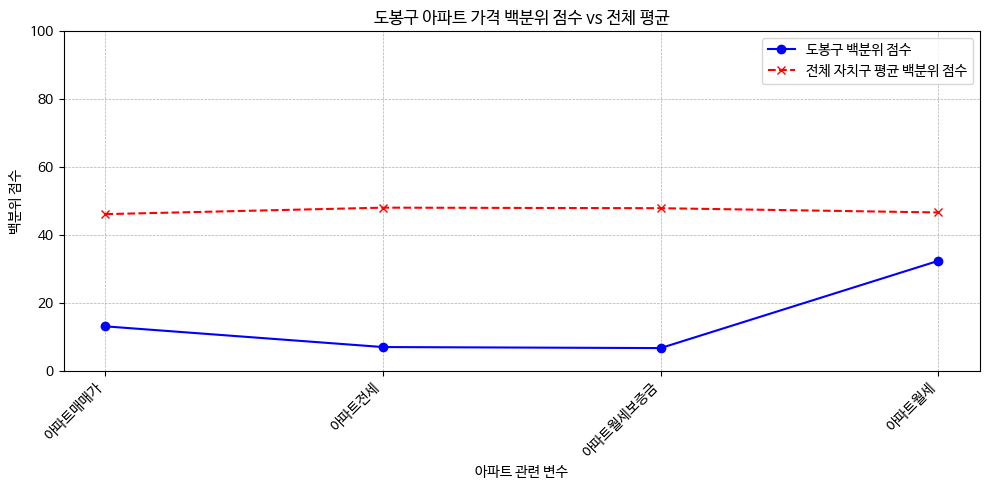
* 노원구

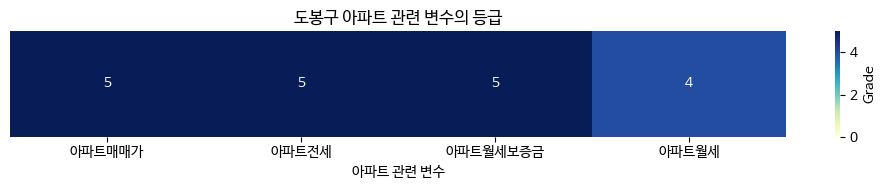




[그래프 1-9] 노원구 아파트 가격 백분위 점수 및 등급

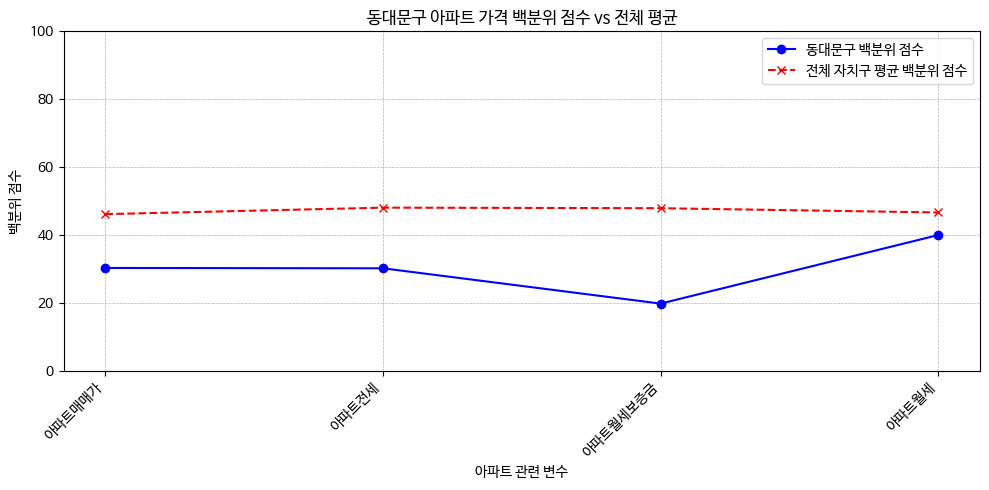
* 도봉구

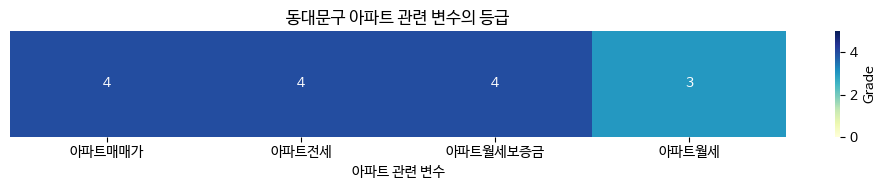




[그래프 1-10] 도봉구 아파트 가격 백분위 점수 및 등급

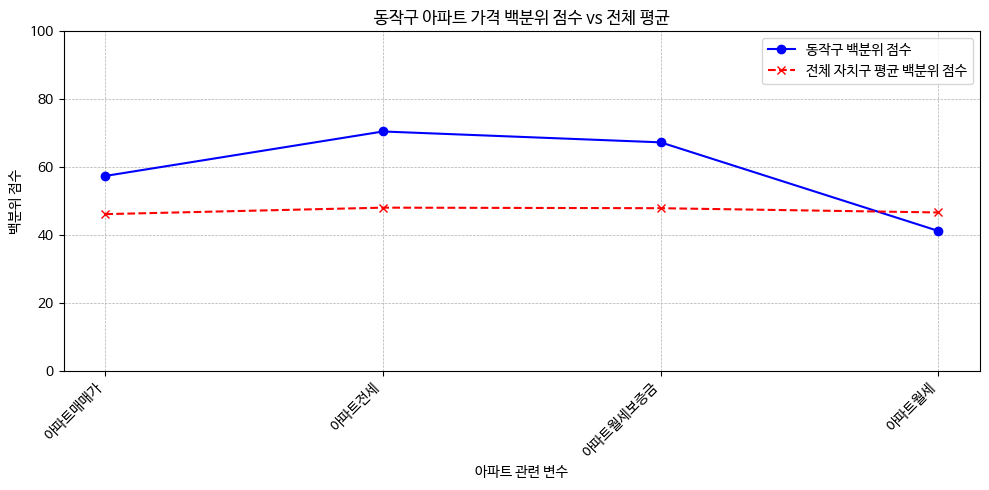
* 동대문구

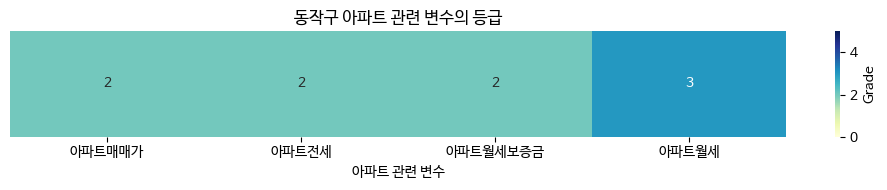




[그래프 1-11] 동대문구 아파트 가격 백분위 점수 및 등급

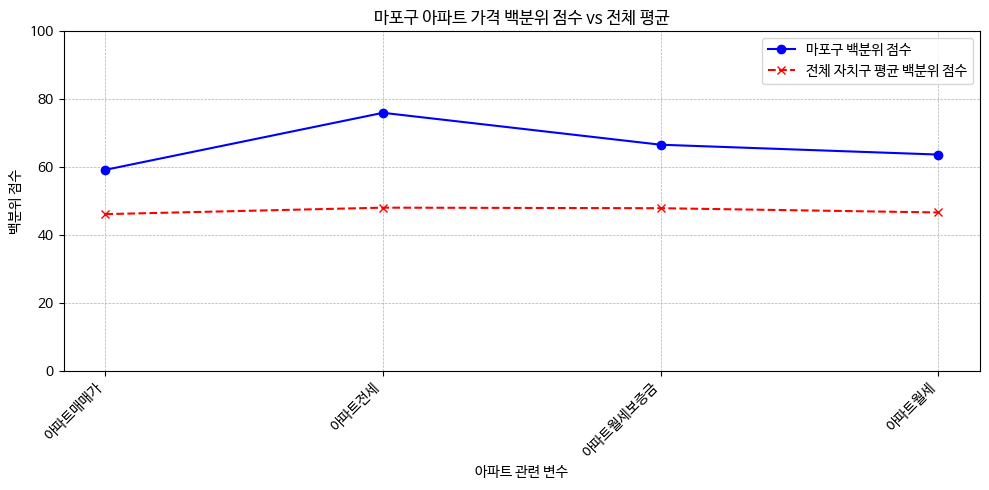
* 동작구

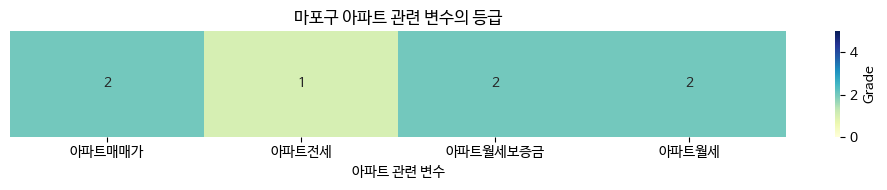




[그래프 1-12] 동작구 아파트 가격 백분위 점수 및 등급

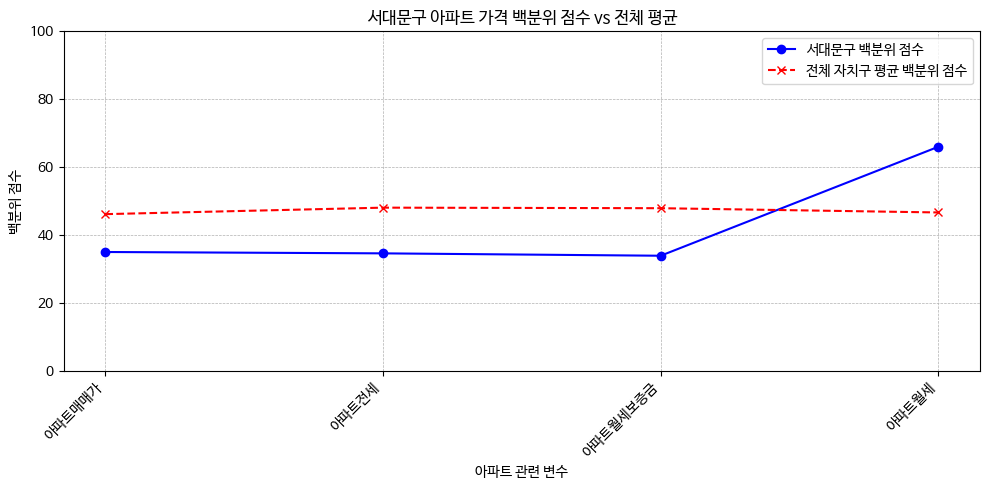
* 마포구

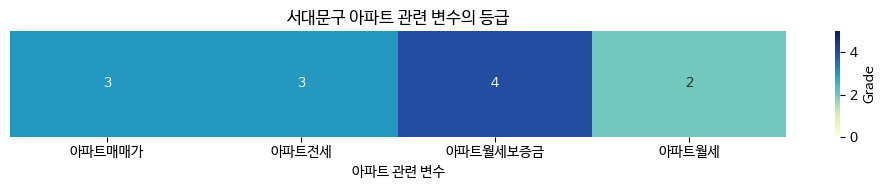




[그래프 1-13] 마포구 아파트 가격 백분위 점수 및 등급

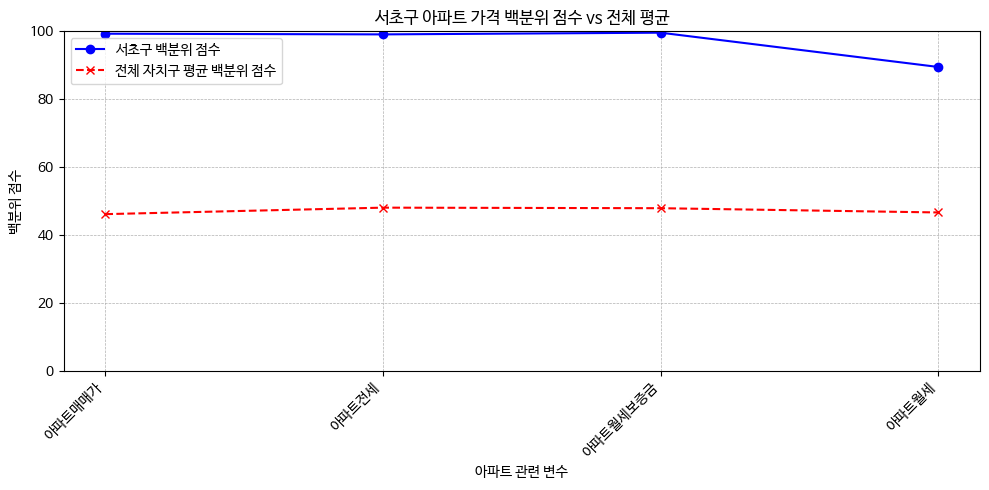
* 서대문구

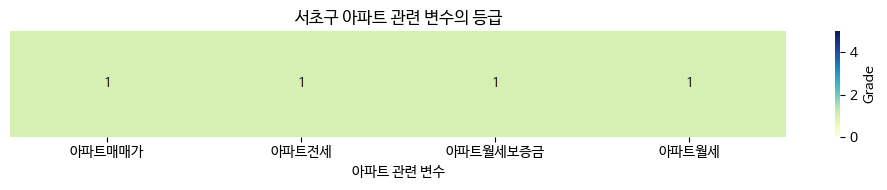




[그래프 1-14] 서대문구 아파트 가격 백분위 점수 및 등급

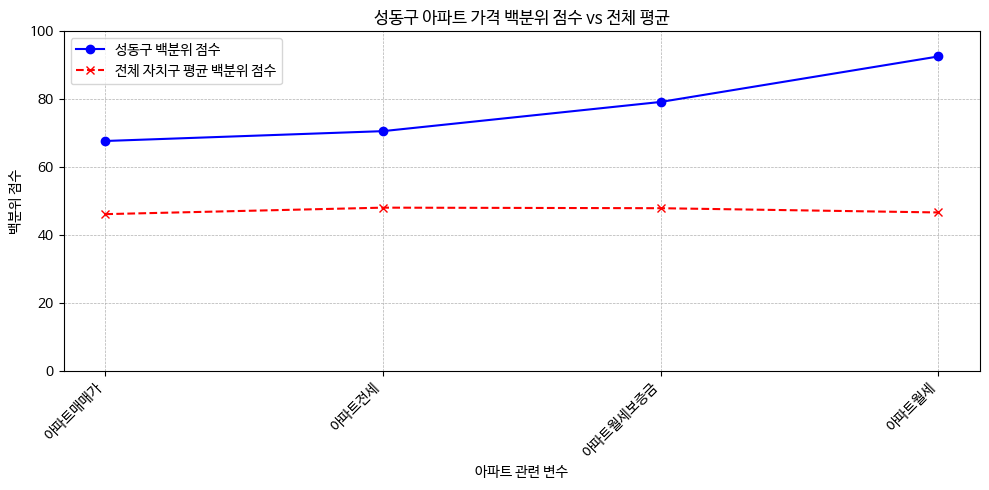
* 서초구

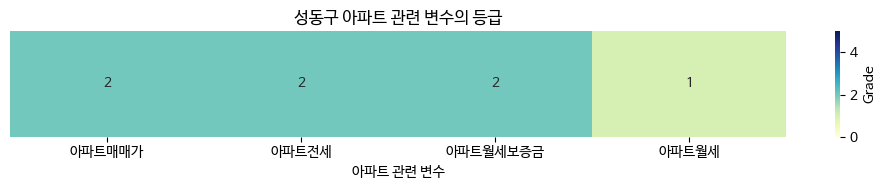




[그래프 1-15] 서초구 아파트 가격 백분위 점수 및 등급

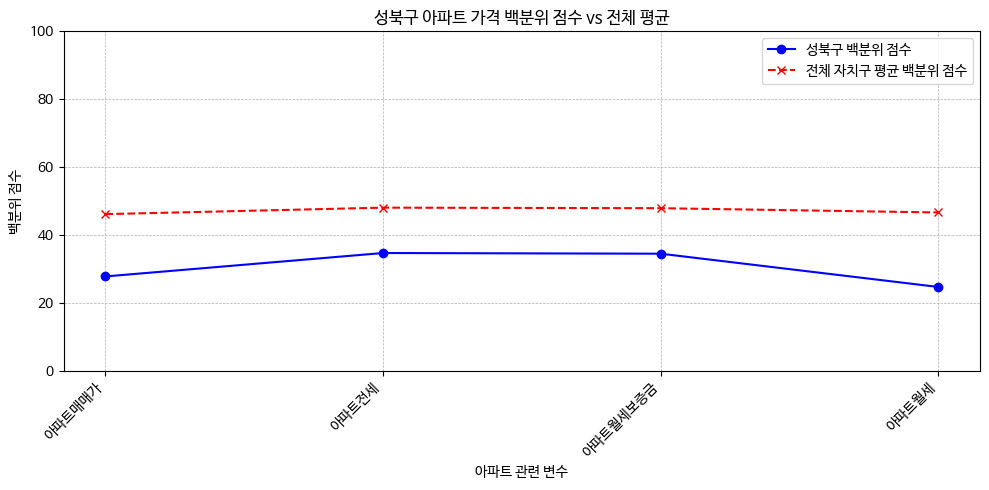
* 성동구

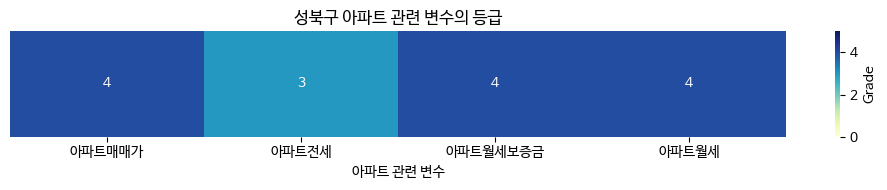




[그래프 1-16] 성동구 아파트 가격 백분위 점수 및 등급

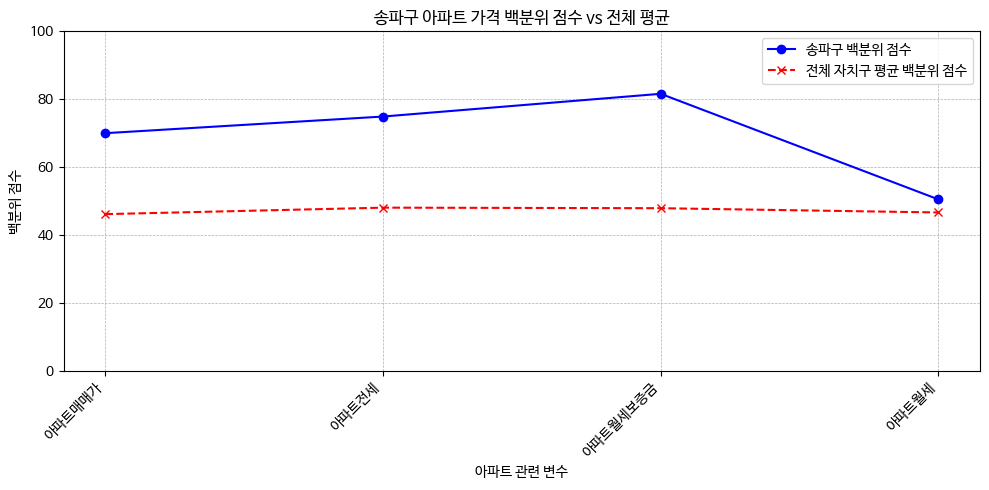
* 성북구

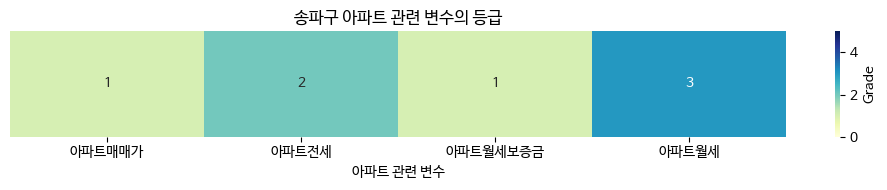




[그래프 1-17] 성북구 아파트 가격 백분위 점수 및 등급

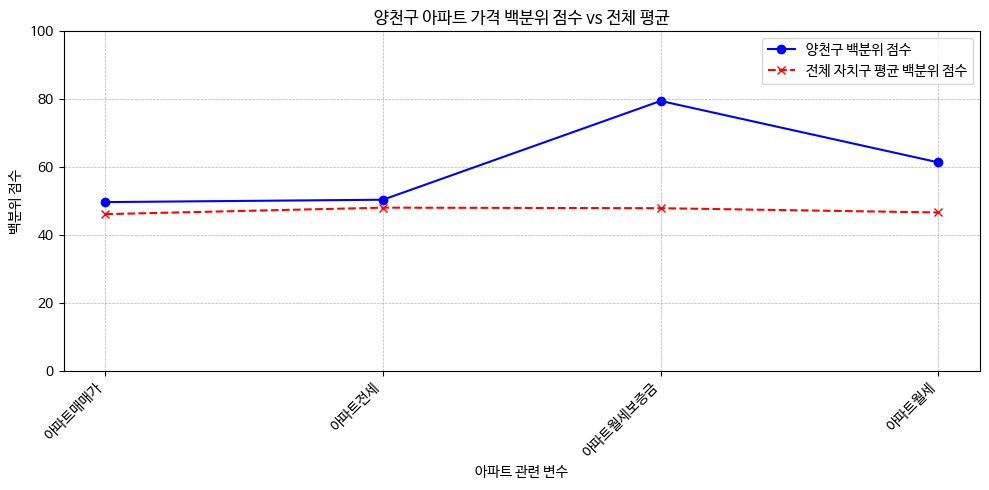
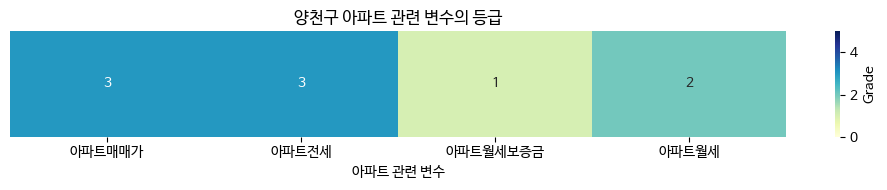
* 송파구





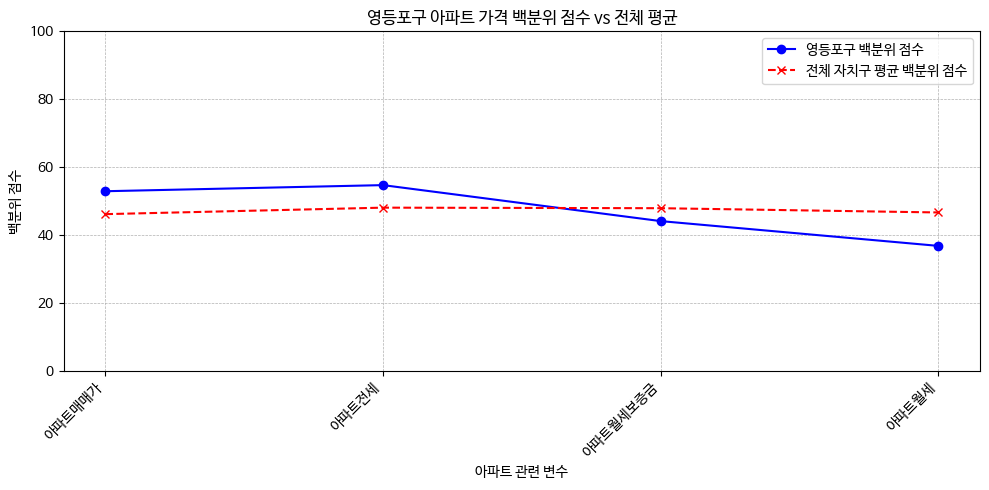
[그래프 1-18] 송파구 아파트 가격 백분위 점수 및 등급

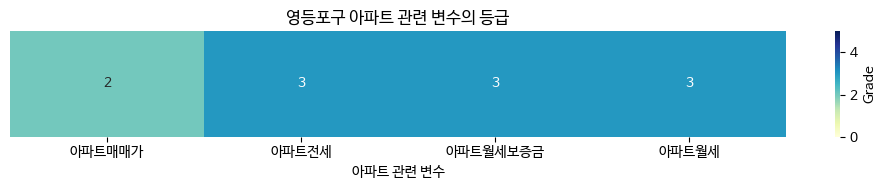
* 양천구

[그래프 1-19] 양천구 아파트 가격 백분위 점수 및 등급

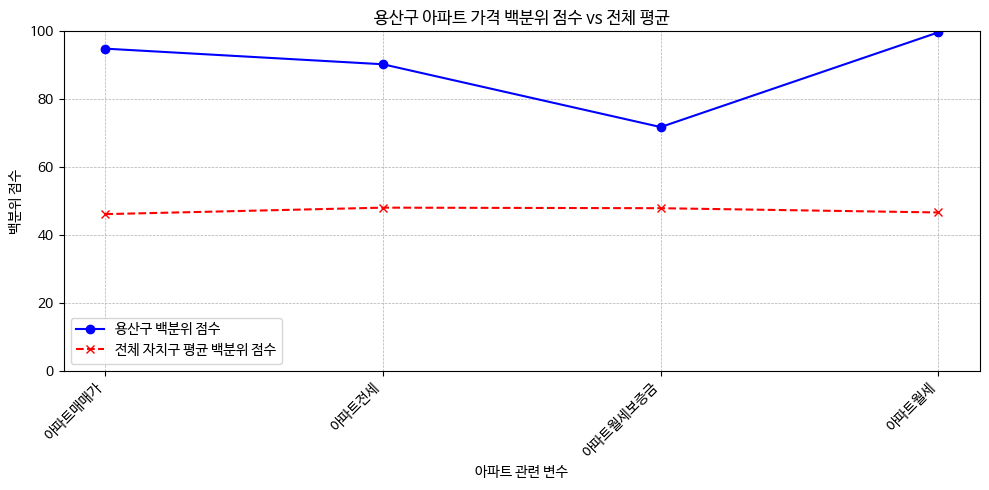
* 영등포구

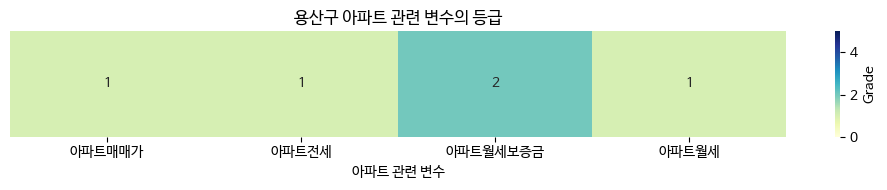




[그래프 1-20] 영등포구 아파트 가격 백분위 점수 및 등급

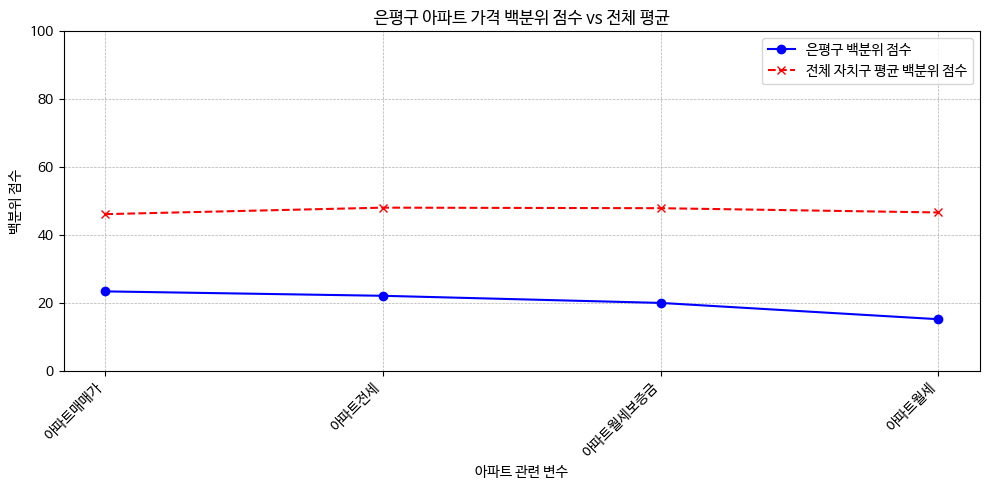
* 용산구

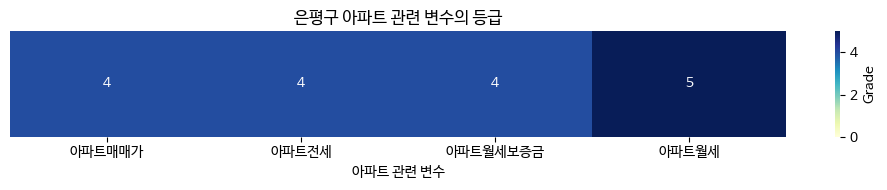




[그래프 1-21] 용산구 아파트 가격 백분위 점수 및 등급

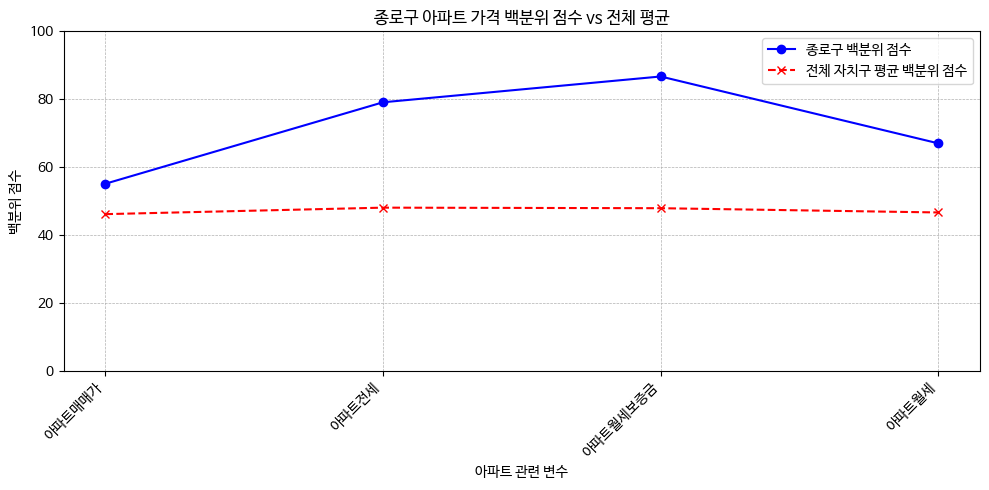
* 은평구

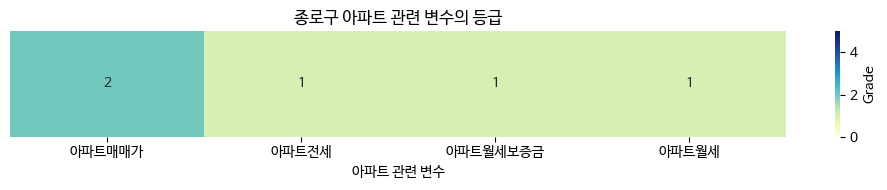




[그래프 1-22] 은평구 아파트 가격 백분위 점수 및 등급

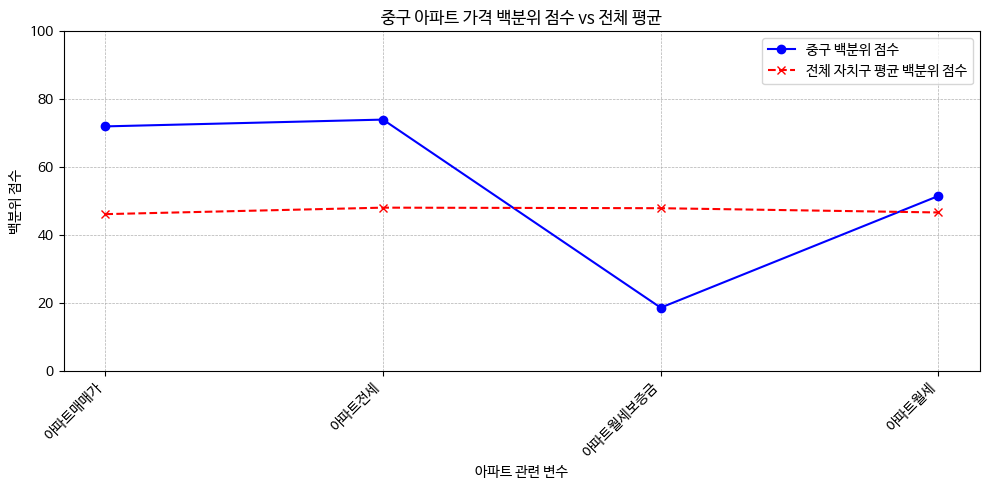
* 종로구

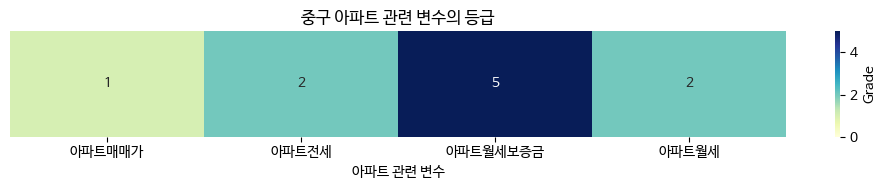




[그래프 1-23] 종로구 아파트 가격 백분위 점수 및 등급

* 중구





[그래프 1-24] 중구 아파트 가격 백분위 점수 및 등급

1. **추천 알고리즘 및 결과**
   1. **추천 알고리즘 개요**

활용 데이터는 3.2.3절에서 통합한 데이터프레임을 사용하였다. 추천 알고리즘은 유저의 선호도와 지역별 가격, 계약 분류, 주거 형태 등 다양한 요소들을 바탕으로 가장 적합한 주거 지역을 추천해주는 알고리즘이다. 여기서 사용된 메인 로직은 유클리디안 거리를 이용한 유사도 기반의 추천 방식이며, 이 외에도 다양한 조건 및 사용자의 선호도를 반영하여 최적화된 추천 결과를 제공한다.

* 1. **주요 과정**
* 유저 정보 설정: 사용자는 계약 분류, 주거 형태, 보증금, 월세, 그리고 선호 순위(3가지)를 입력한다.
* 유클리디안 거리 계산: 월세인 경우, 보증금 및 월세의 단위를 통일 후 사용자가 입력한 값과 각 지역의 값을 바탕으로 유클리디안 거리를 계산한다.
* 선호도 반영 정렬: 입력된 선호 순위를 바탕으로, 해당 순위에 따라 지역을 정렬한다. 예를 들어, 사용자가 인구 전입량을 1순위로 선정했을 경우, 인구 전입량이 높은 지역을 우선적으로 추천한다.
* 추천 결과 제공: 위의 과정을 통해 나온 최종 지역 리스트를 사용자에게 제공한다. 이때, 결과는 JSON 형태로 반환되어, 다양한 플랫폼에서의 활용이 용이하다.
  1. **추천 알고리즘 결과**
     1. **샘플 데이터 셋**
* 사용자의 정보(user name)
* user는 주택 계약을 "전세" 형태를 선택
* 원하는 주거형태는 오피스텔이며, 전세금으로 지불 가능한 금액은 10,000(만 원)을 입력
* 전세 계약을 희망하므로, 월세는 지불하지 않음(0으로 입력).
* 해당 사용자는 지역 선택 시 인구 유입량, 대형마트 수, 그리고 지하철역 수 순으로 중요하게 생각한다.

|  | **사용자** | **계약 분류** | **주거형태** | **보증금(전세금)** | **월세** | **선호1순위** | **선호2순위** | **선호3순위** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | user | 전세 | 오피스텔 | 10000 | 0 | 인구 | 대형마트 | 지하철 |

**[표 2-1] 추천 알고리즘 샘플 데이터 셋**

* + 1. **추천 결과**
* 위 사용자의 선호도와 예산을 기반으로 추천된 지역 목록이다.
* 각 지역은 해당 지역의 오피스텔 전세 평균가, 인구 유입량, 대형마트 수, 그리고 지하철역 수에 대한 정보를 포함하고 있다.
* 첫 번째 추천 결과를 확인해보면, 강서구 화곡동은 오피스텔 전세 평균가가 10,320.967740 (만 원)으로 사용자의 예산과 가까우며, 인구 유입량이 9313.0, 대형마트 수가 6, 지하철역 수가 4이다. 이러한 정보를 통해 해당 지역이 사용자의 선호도와 잘 맞는지를 판단할 수 있다.

|  | **자치구** | **행정동** | **오피스텔전세** | **인구 유입량** | **대형마트 수** | **지하철역 수** |
| --- | --- | --- | --- | --- | --- | --- |
| **0** | 강서구 | 화곡동 | 10320.967740 | 9313.0 | 6 | 4 |
| **1** | 구로구 | 구로동 | 10190.116280 | 6323.0 | 12 | 5 |
| **2** | 광진구 | 화양동 | 10082.976740 | 4925.0 | 2 | 3 |
| **3** | 동작구 | 사당동 | 9650.000000 | 4764.0 | 3 | 14 |
| **4** | 노원구 | 공릉동 | 10250.000000 | 3574.0 | 3 | 4 |
| **5** | 서대문구 | 연희동 | 9612.727273 | 2664.0 | 1 | 0 |
| **6** | 구로구 | 오류동 | 10431.041670 | 2350.0 | 2 | 3 |
| **7** | 성동구 | 성수1가동 | 9818.181818 | 1004.0 | 1 | 2 |
| **8** | 은평구 | 신사동 | 9561.538462 | 960.0 | 1 | 2 |
| **9** | 서대문구 | 북아현동 | 10309.523810 | 357.0 | 1 | 0 |
| **10** | 중구 | 필동 | 9993.181818 | 343.0 | 0 | 3 |
| **11** | 서초구 | 내곡동 | 9760.000000 | 279.0 | 0 | 1 |

**[표 2-2] 추천 알고리즘 샘플 데이터 셋 결과**

1. **활용방안**
   1. **부동산 중개업 및 플랫폼 활용**

* **고객 맞춤형 서비스 제공**

중개업자나 부동산 플랫폼은 주어진 데이터와 알고리즘을 활용하여 고객의 주거 선호도와 경제적 상황에 딱 맞는 주거지역을 추천함으로써 맞춤형 서비스를 제공할 수 있다. 이를 통해 고객의 탐색 시간을 절약하고, 중개의 효율성을 향상시킬 수 있다.

* **경쟁력 강화**

다양한 부동산 중개 플랫폼 중에서도 해당 알고리즘을 활용한 서비스는 고객에게 더 정확하고 빠른 매칭을 제공함으로써 경쟁력을 강화할 수 있다

* 1. **정책 제안 및 도시 계획:**
* **주거 환경 개선**

지역별 인구 유입량, 대형마트 수, 지하철역 수 등의 데이터는 지자체나 정부의 주거 환경 개선 정책의 방향성을 설정하는 데 근거 자료로 사용될 수 있다. 특히, 높은 인구 유입량을 보이는 지역에는 추가적인 교통시설이나 편의시설 투자가 필요할 수 있다.

* **주거지역 발전 전략 수립**

해당 데이터를 활용하여 지역별 주거지역 발전 전략을 수립하며, 주민들의 삶의 질을 향상시키는 방향으로 계획을 세울 수 있다.

* 1. **주민들의 의사결정 지원:**
* **주거지 선택의 근거 마련**

주민들은 제시된 데이터를 통해 주거지 선택 시 지역별 편의시설 및 교통 인프라 등을 비교 분석하고, 이를 바탕으로 최적의 주거 환경을 선택할 수 있다.

* **생활 편의성 판단**

주민들은 대형마트 수, 지하철역 수 등의 데이터를 활용하여 해당 지역의 생활 편의성을 판단하고, 이를 바탕으로 주거 이전 여부나 장기 거주 여부를 결정할 수 있다.

* 1. **데이터 기반 연구 및 학술 활용:**
* **도시 및 주거 문화 연구**

학계 연구자들은 해당 데이터와 알고리즘 결과를 활용하여 도시나 주거 문화에 관한 심층 연구를 수행할 수 있다. 특히, 도시 내 인구 이동 및 주거 선호도 변화와 같은 동향 분석이 가능하다.

* **정책 제안 연구**

주거 문제나 도시 문제에 관한 학술 연구를 통해 보다 실용적이고 효과적인 정책 제안이 가능하며, 해당 데이터는 그러한 연구의 근거 자료로 활용될 수 있다.

1. **참고문헌**

서울연구원, “서울시 1인가구 실태조사 및 제도개선 연구용역 보고서”, 2022. 02

통계청 사회통계국 사회통계기획과, “2022 통계로 보는 1인가구”, 통계청 보도자료, 2022.12.07

이신애, 황원경, 정승환, “2022년 한국1인가구 보고서”, KB금융지주 경영연구소, 2022.10