

인천광역시 아파트 실거래가 분석을 통한 부동산 시장 동 향 시각화

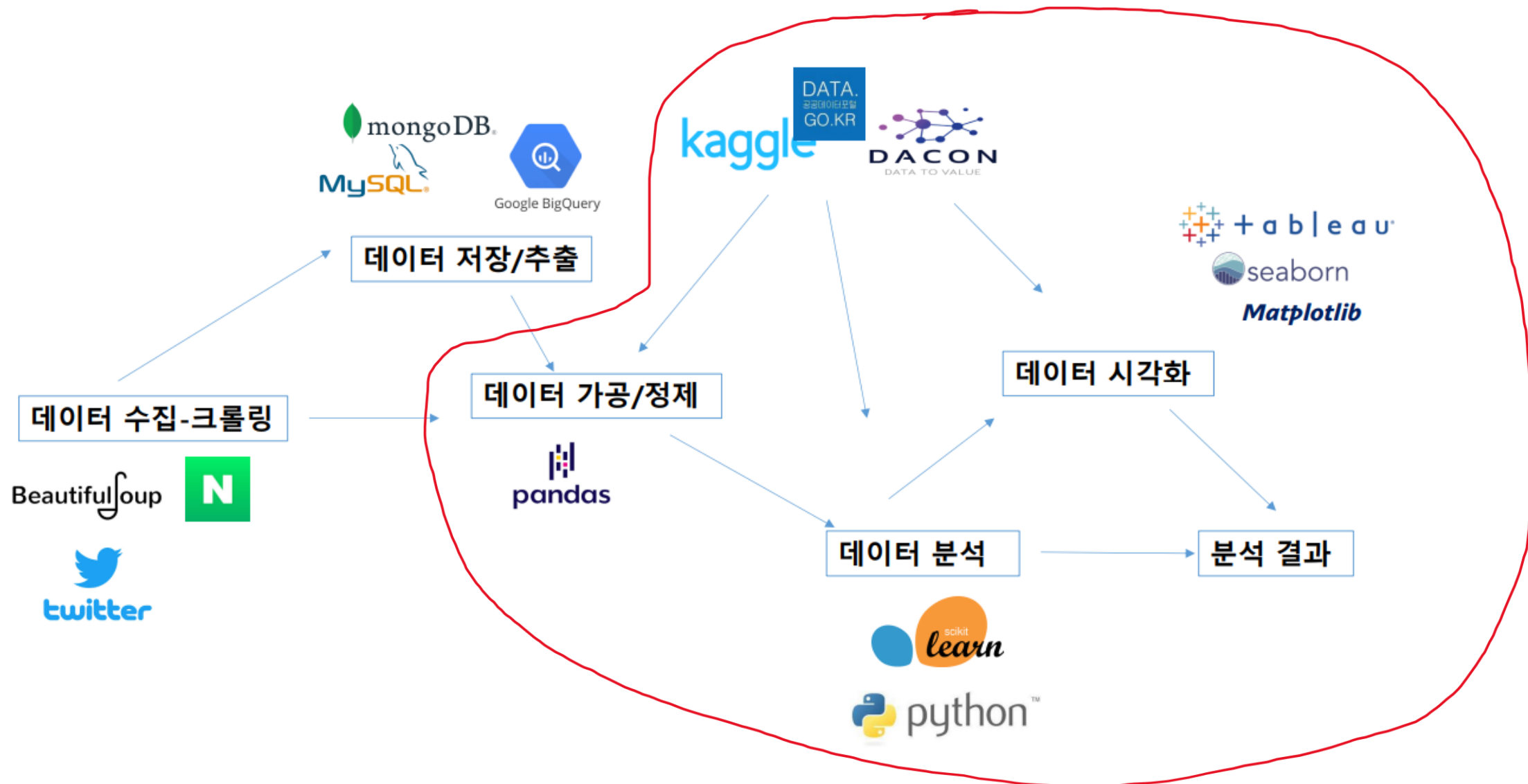
202044001 조찬희

2025.12.07

프로젝트 개요

본 프로젝트는 공공데이터포털에서 제공하는 인천광역시 아파트 실거래가 데이터를 활용하여, 2025년 최근 데이터를 기준으로 인천 부동산 시장의 현황을 파악하고 지역별 특성을 시각적으로 분석하는 것을 목표로 합니다.

프로젝트 범위



데이터 출처

- 데이터: 국토교통부 실거래가 공개시스템 (아파트 매매 실거래 상세 자료)
- 수집 방법: 국토교통부 사이트에서 인천광역시 전체 지역을 대상으로 한 최근 1년의 데이터를 CSV 파일 형태로 다운로드하여 사용
- 규모: 27763건

데이터 로드 및 전처리

```
13 # =====
14 # 1. 데이터 로드 및 전처리 (Preprocessing)
15 # =====
16 file_path = '아파트(매매)_실거래가_20251207162315.csv'
17 # 국토부 파일은 보통 헤더(설명)가 15줄 정도 있어서 header=15 옵션이 필요합니다.
18 df = pd.read_csv(file_path, encoding='cp949', header=15)
19
20 # (1) 숫자 변환: "15,000" -> 15000
21 df['거래금액'] = df['거래금액(만원)'].str.replace(',', '', True).astype(int)
22
23 # (2) 파생 변수 생성
24 df['구'] = df['시군구'].apply(lambda x: x.split()[1]) # 구 이름 추출
25 df['평당가'] = df['거래금액'] / (df['전용면적(m²)'] / 3.305785) # 평당가 계산
26 df['건축년대'] = (df['건축년도'] // 10) * 10 # 10년 단위 연식 그룹화 (예: 1990, 2000)
```

데이터 분석 1 - 지역별 평균금액과 거래량

```
28 # =====
29 # 2. 데이터 분석 (Data Analysis) - 수치 확인
30 # =====
31 # [분석 1] 지역별 시장 현황 요약
32 district_stats = df.groupby('구').agg(
33     평균거래금액=('거래금액', 'mean'),
34     평균평당가=('평당가', 'mean'),
35     거래량=('거래금액', 'count')
36 ).sort_values(by='평균거래금액', ascending=False)
37
38 print("=== [1] 지역별 아파트 시장 현황 (비싼 순) ===")
39 print(district_stats)
40 print("\n")
```

```
=== [1] 지역별 아파트 시장 현황 (비싼 순) ===
      평균거래금액  평균평당가  거래량
구
연수구    60049.6  2367.7    5020
서구      44500.1  1843.2    6135
부평구    37240.5  1788.4    4331
남동구    34949.7  1548.0    4218
중구      33885.6  1446.3    1528
미추홀구  32491.3  1497.1    3063
계양구    27629.2  1402.2    2694
동구      23207.6  1226.0     604
강화군    22086.0  1061.6     170
```

데이터 분석 2 - 평당가 상위 아파트

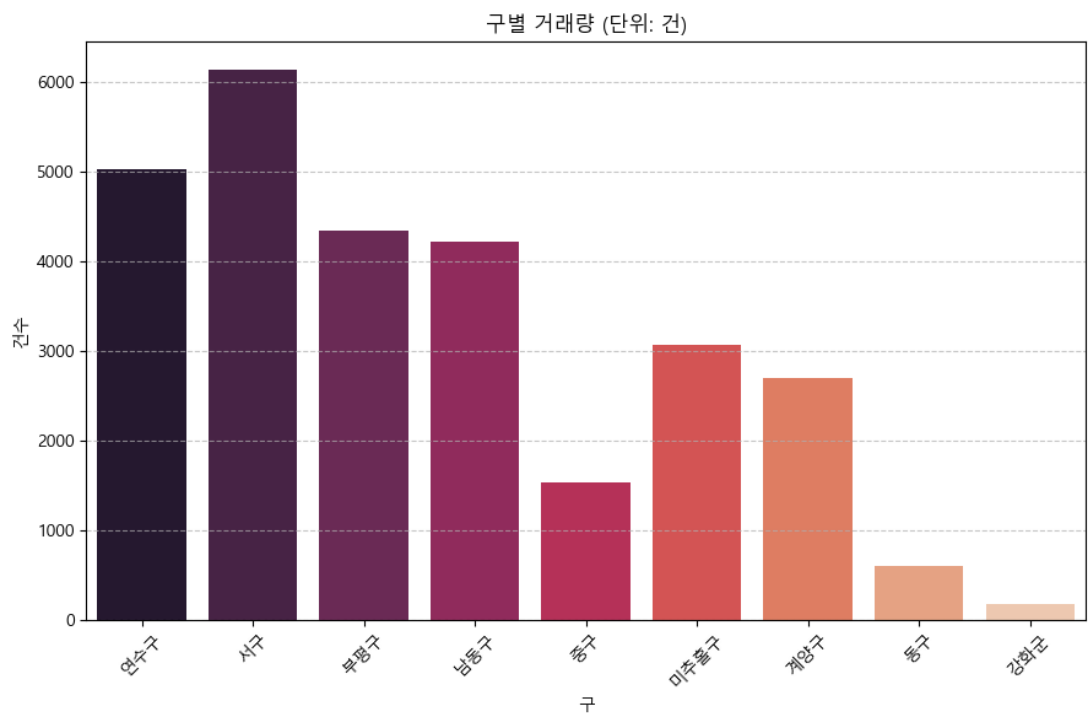
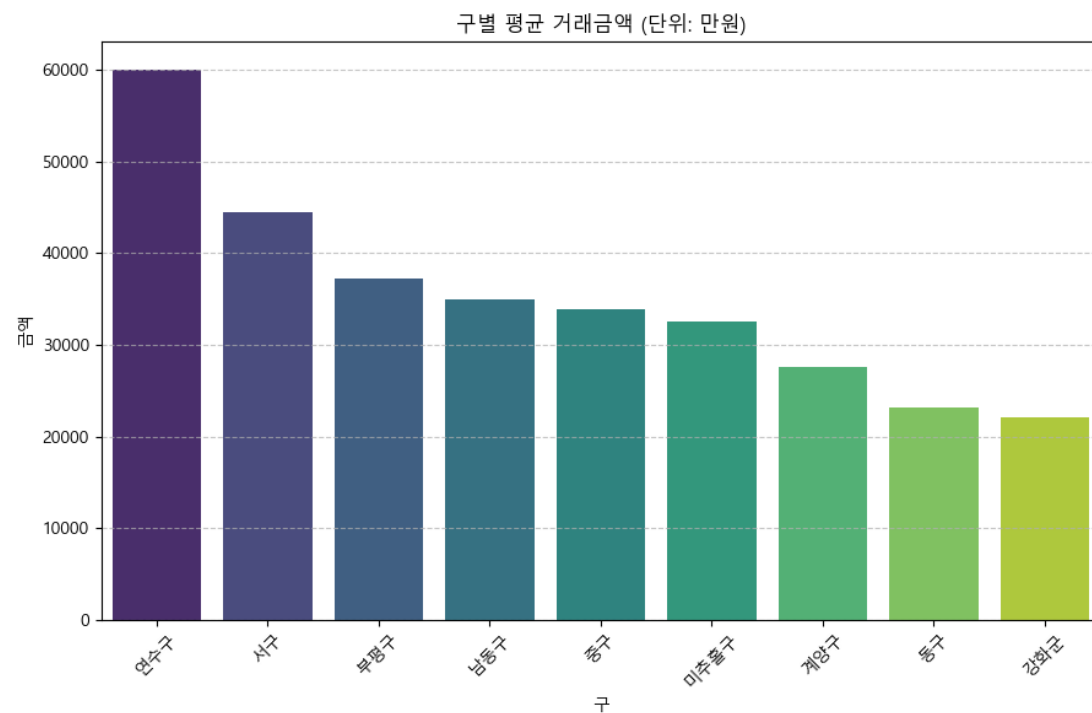
```
42 # [분석 2] 평당가 기준 상위 10개 아파트
43 top_apartments = df.sort_values(by='평당가', ascending=False)[
44     ['단지명', '구', '전용면적(㎡)', '건축년도', '거래금액', '평당가']
45 ].head(10)
46
47 print("=== [2] 평당가 기준 TOP 10 아파트 ===")
48 print(top_apartments)
49 print("\n")
```

```
=== [2] 평당가 기준 TOP 10 아파트 ===
   단지명  구  전용면적(㎡)  건축년도  거래금액  평당가
20933  송도더샵파크애비뉴  연수구    84.4   2018   155000  6071.1
26665  더샵송도센트럴파크Ⅲ  연수구   119.8   2022   202000  5575.4
10620  더샵송도센트럴파크Ⅲ  연수구   110.8   2022   185500  5532.0
3321   송도더샵파크애비뉴  연수구    68.6   2018   110000  5297.7
3359   더샵송도센트럴파크Ⅲ  연수구   102.5   2022   162800  5251.1
22524  더샵송도센트럴파크Ⅲ  연수구   102.5   2022   161000  5193.0
22523  더샵송도센트럴파크Ⅲ  연수구   102.5   2022   161000  5193.0
825    송도더샵파크애비뉴  연수구    84.0   2018   131000  5152.4
1906   송도더샵파크애비뉴  연수구    68.6   2018   106800  5143.6
4441   송도더샵파크애비뉴  연수구    68.6   2018   105500  5081.0
```

데이터 시각화 1 – 지역별 거래금액과 거래량 시각화

```
51 # =====
52 # 3. 데이터 시각화 (Visualization) - 그래프 확인
53 # =====
54
55 # (1) [Bar Chart] 구별 평균 거래금액 & 거래량 (subplot으로 한 번에 출력)
56 fig, ax = plt.subplots( nrows= 1, ncols= 2, figsize=(18, 6))
57
58 # 왼쪽: 평균 거래금액
59 sns.barplot(x=district_stats.index, y=district_stats['평균거래금액'], palette='viridis', ax=ax[0])
60 ax[0].set_title('구별 평균 거래금액 (단위: 만원)')
61 ax[0].set_ylabel('금액')
62 ax[0].tick_params(axis='x', rotation=45)
63 ax[0].grid(axis='y', linestyle='--', alpha=0.7)
64
65 # 오른쪽: 거래량
66 sns.barplot(x=district_stats.index, y=district_stats['거래량'], palette='rocket', ax=ax[1])
67 ax[1].set_title('구별 거래량 (단위: 건)')
68 ax[1].set_ylabel('건수')
69 ax[1].tick_params(axis='x', rotation=45)
70 ax[1].grid(axis='y', linestyle='--', alpha=0.7)
71
72 plt.tight_layout()
73 plt.show()
```

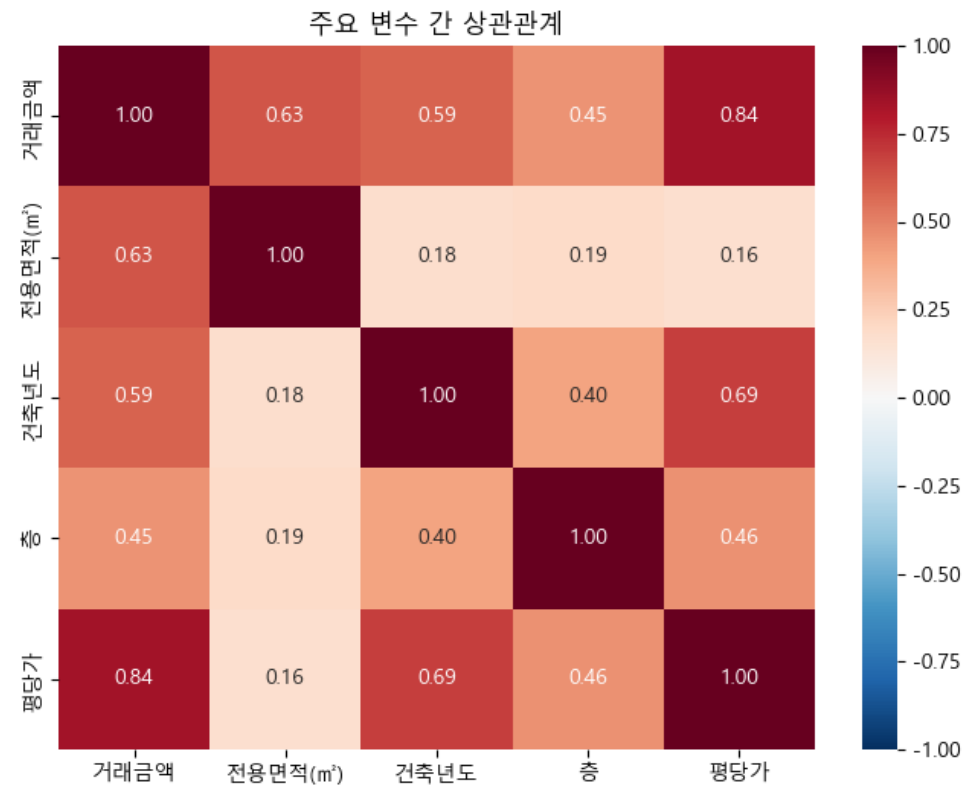

데이터 시각화 1 - 지역별 거래금액과 거래량 시각화



데이터 시각화 2 – 주요변수 간 상관관계

```
75 # (2) [Correlation Heatmap] 상관관계 분석
76 numeric_cols = ['거래금액', '전용면적(m²)', '건축년도', '층', '평당가']
77 corr_matrix = df[numeric_cols].corr()
78
79 plt.figure(figsize=(8, 6))
80 sns.heatmap(corr_matrix, annot=True, fmt='.2f', cmap='RdBu_r', vmin=-1, vmax=1)
81 plt.title('주요 변수 간 상관관계')
82 plt.show()
```

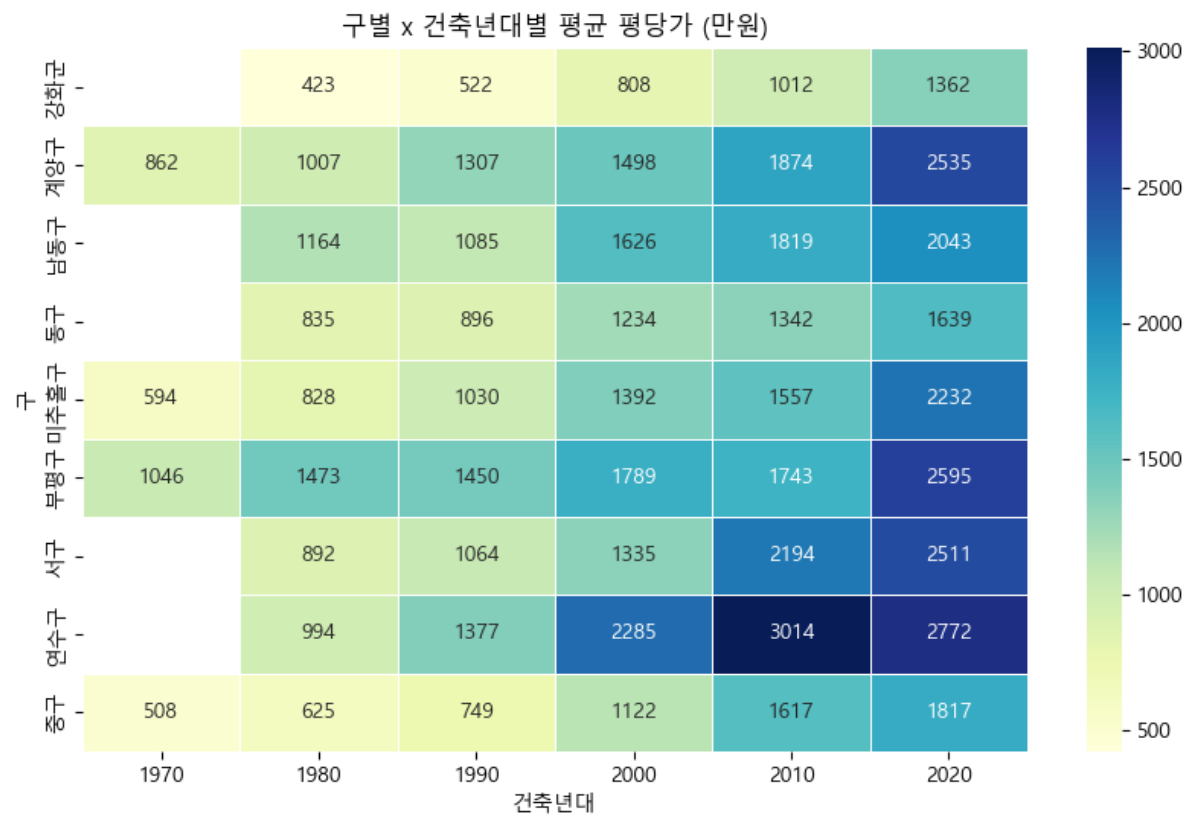
데이터 시각화 2 – 주요변수 간 상관관계



데이터 시각화 3 - 건축년대별 평당가 시각화

```
84 # (3) [Pivot Heatmap] 구별 x 건축년대별 평당가 (핵심 분석)
85 pivot_price = df.pivot_table(index='구', columns='건축년대', values='평당가', aggfunc='mean')
86
87 plt.figure(figsize=(10, 6))
88 sns.heatmap(pivot_price, annot=True, fmt='.0f', cmap='YlGnBu', linewidths=0.5)
89 plt.title('구별 x 건축년대별 평균 평당가 (만원)')
90 plt.xlabel('건축년대')
91 plt.ylabel('구')
92 plt.show()
```

데이터 시각화 3 – 건축년대별 평당가 시각화



결론

- 신도시 지역인 연수구, 서구가 평균 거래금액과 거래량이 각각 가장 높다.
 - 평균 거래금액과 거래량은 같이 높은 경향을 보인다.
 - 모든 구에서 건축연도가 최근일수록 평균 평당가가 높아지는 경향이 있다.
- ⇒ 거래금액 상위 지역일수록 거래도 활발한 편이며, 지역에 상관없이 신축일수록 평당가가 확실히 높다.

감사합니다