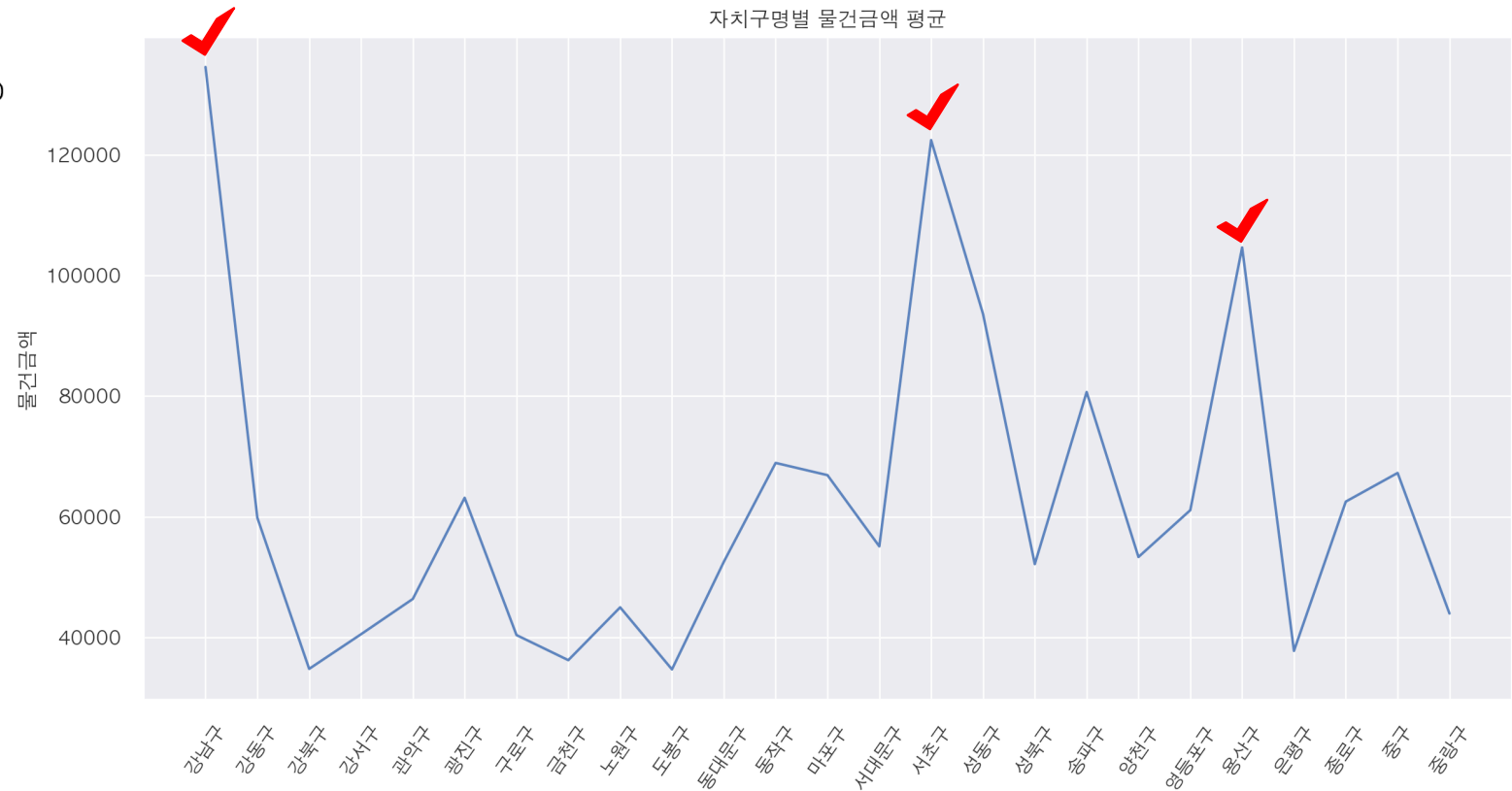


# 3 데이터 시각화 및 분석 : 자치구명 vs 물건금액

```
# 1) 자치구명 vs 물건금액
group_gu = df.groupby('자치구명')
group_gu_price = group_gu['물건금액'].mean()

plt.figure(figsize=(20, 7))
sns.lineplot(data=group_gu_price)
plt.xticks(rotation=55)
plt.title('자치구명별 물건금액 평균')
plt.show()
```

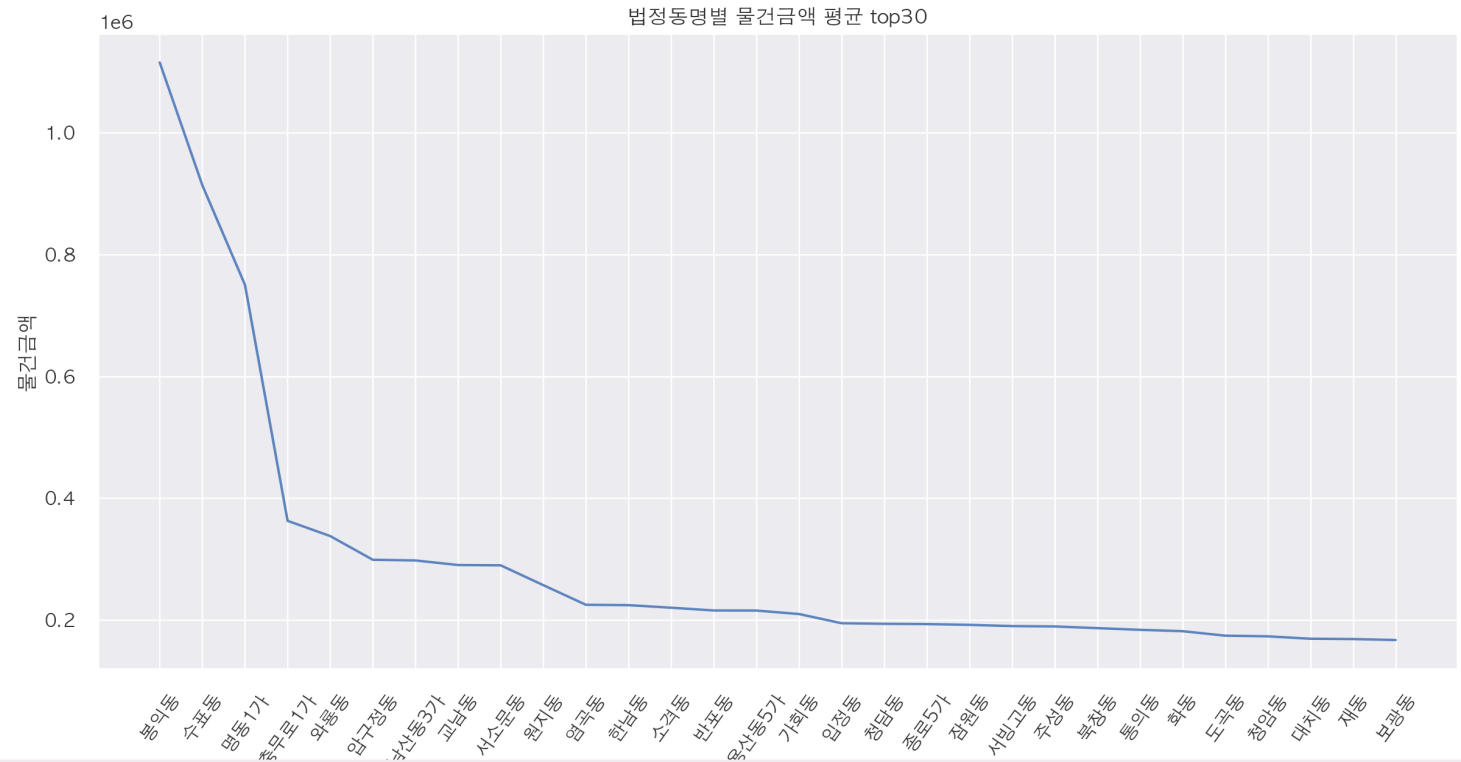


자치구별 부동산 거래가격 top3: 강남구, 서초구, 용산구

# 3 데이터 시각화 및 분석 : 법정동명 vs 물건금액

```
# 2) 법정동명 vs 물건금액
group_dong = df.groupby('법정동명')
group_dong_price = group_dong['물건금액'].mean()
group_dong_price = group_dong_price.sort_values

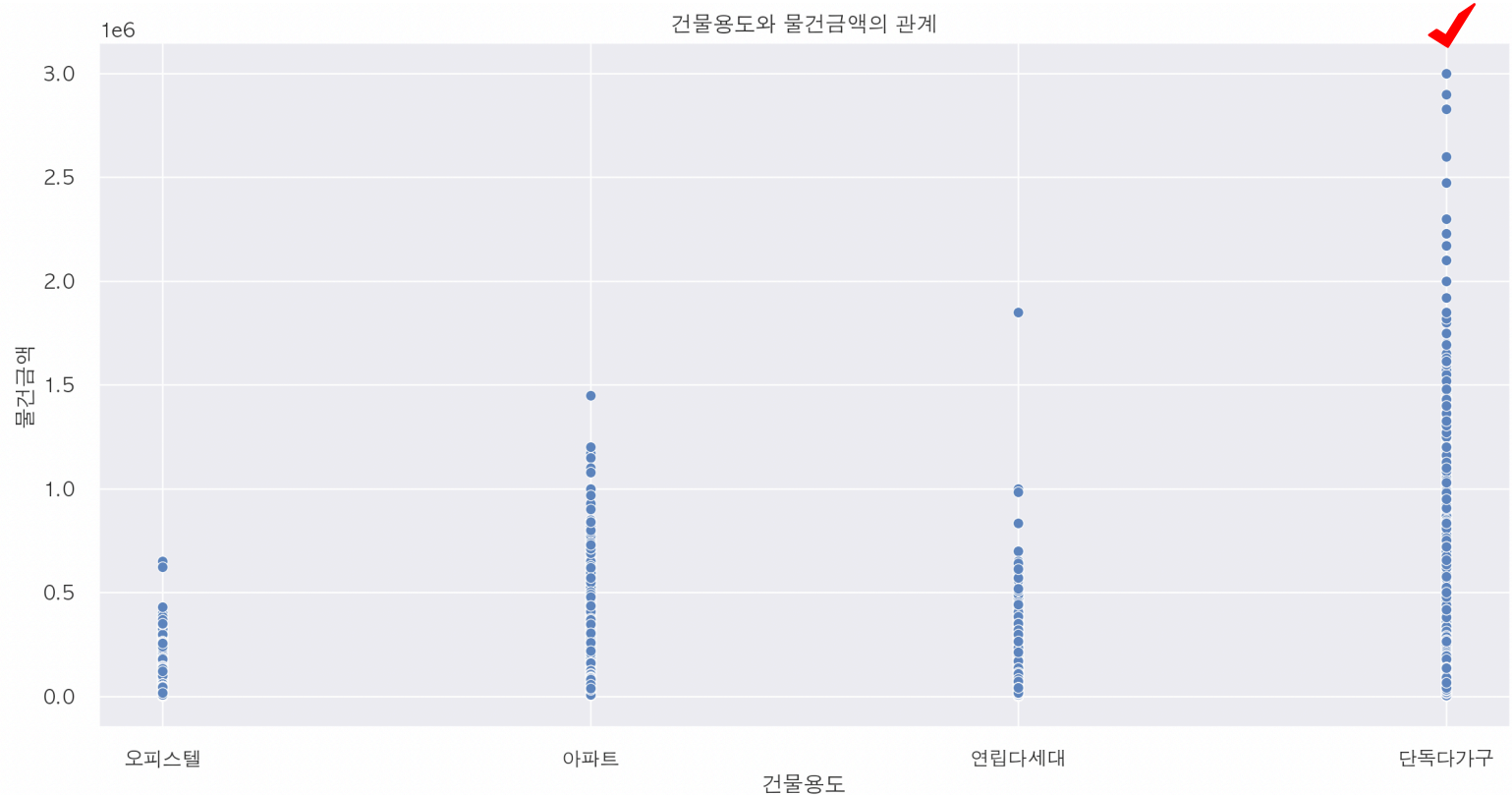
plt.figure(figsize=(20, 7))
sns.lineplot(data=group_dong_price[:30])
plt.xticks(rotation=55)
plt.title('법정동명별 물건금액 평균 top30')
plt.show()
```



봉익동(종로구), 수표동(중구), 명동1가(중구) 순으로 거래가격이 높고 주로 단독다가구(상가) 건물임

# 3 데이터 시각화 및 분석 : 건물용도 vs 물건금액

```
# 3) 건물용도 vs 물건금액
plt.figure(figsize=(20, 7))
sns.scatterplot(data=df, x="건물용도", y="물건금액")
plt.title("건물용도와 물건금액의 관계")
plt.show()
df['건물용도'].value_counts()
'''
아파트      290943
연립다세대   229303
오피스텔      62717
단독다가구    48542
Name: 건물용도, dtype: int64
'''
```



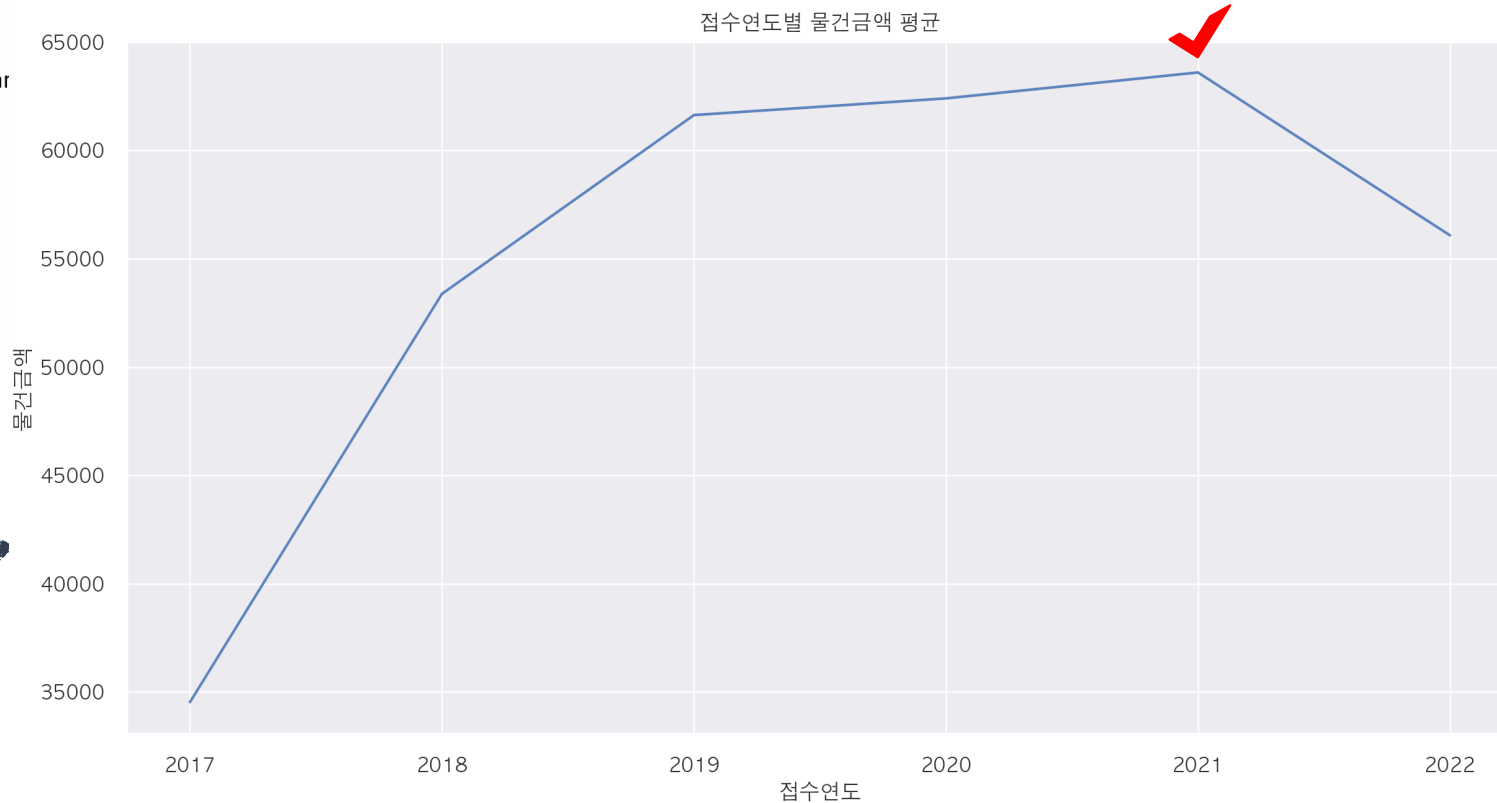
단독다가구(상가) 개수는 다른 용도에 비하면 적지만, 거래가격은 상대적으로 높은 편

# 3 데이터 시각화 및 분석 : 접수연도 vs 물건금액

# 4) 접수연도 vs 물건금액

```
group_year1 = df.groupby('접수연도')  
group_year1_price = group_year1['물건금액'].mean
```

```
plt.figure(figsize=(20, 7))  
sns.lineplot(data=group_year1_price)  
plt.xlabel('접수연도')  
plt.ylabel('물건금액')  
plt.title('접수연도별 물건금액 평균')  
plt.show()
```



2017년부터 2021년까지 부동산 거래가격이 상승하다가 2021년부터 현재까지 떨어지고 있는 추세

# 3 데이터 시각화 및 분석 : 건물면적 vs 물건금액

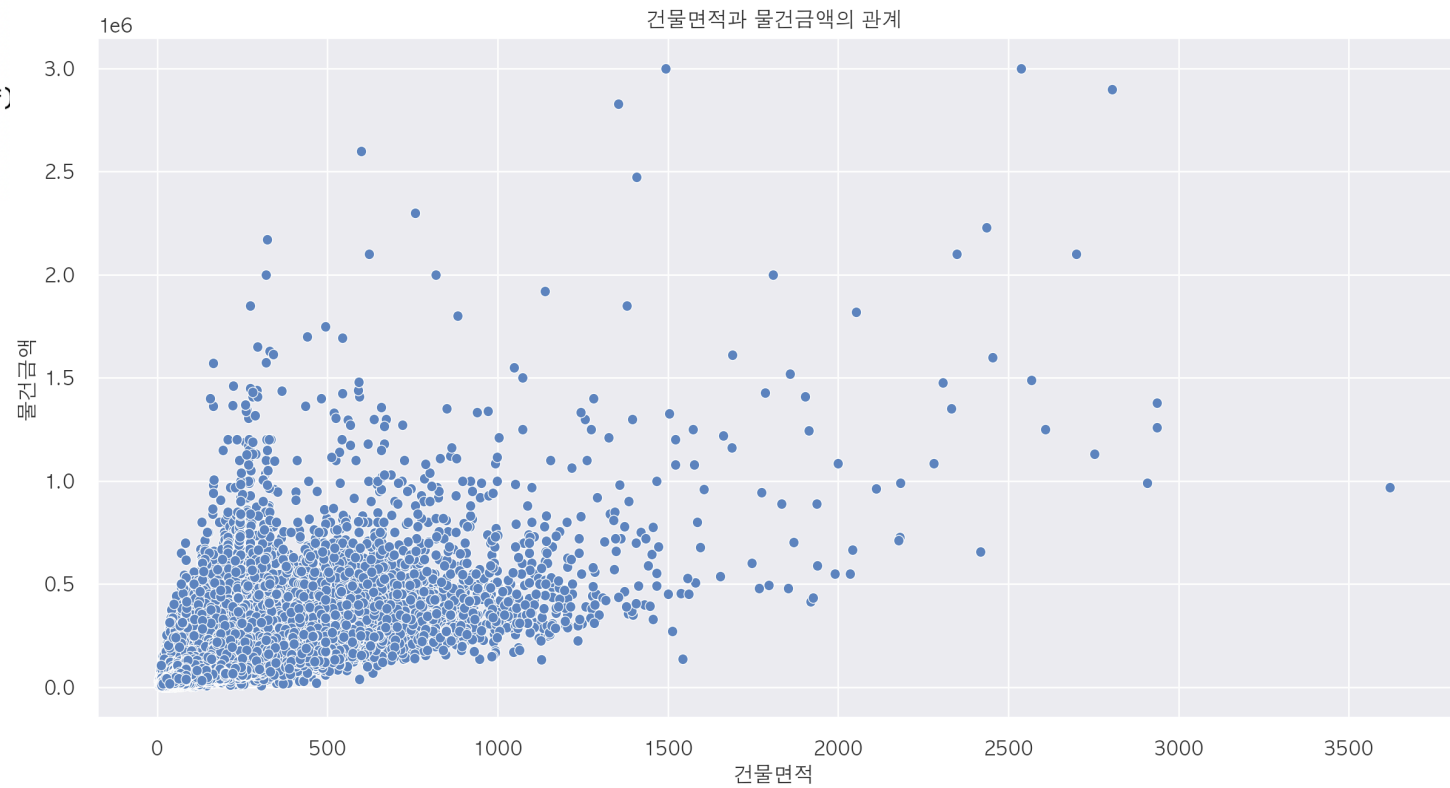
```
# 5) 건물면적 vs 물건금액
```

```
plt.figure(figsize=(20,7))
```

```
sns.scatterplot(x='건물면적', y='물건금액', data=df)
```

```
plt.title("건물면적과 물건금액의 관계")
```

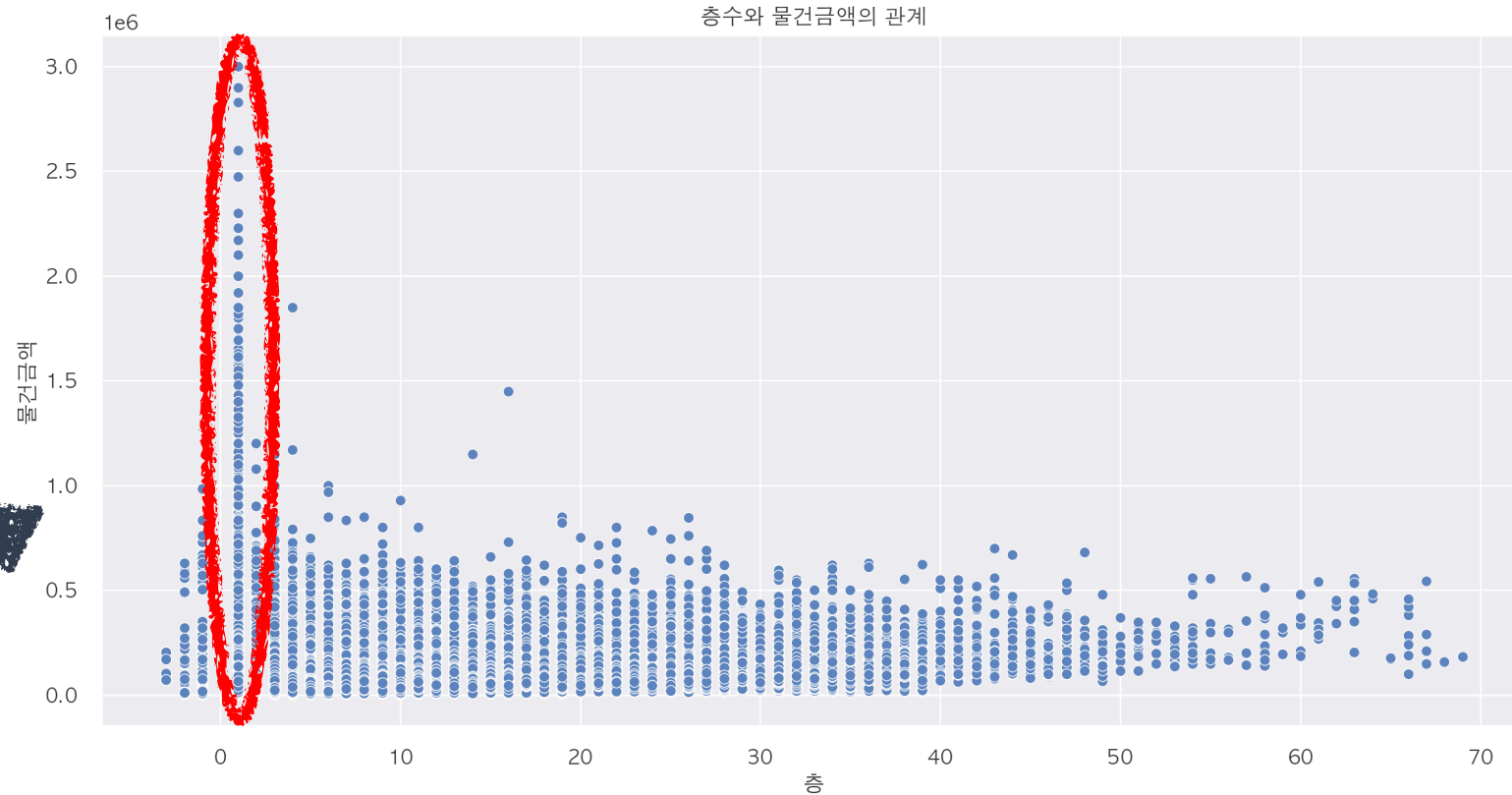
```
plt.show()
```



건물 면적과 부동산 거래가격 간의 유의미한 상관성이 없음

# 3 데이터 시각화 및 분석 : 층 vs 물건금액

```
# 6) 층 vs 물건금액
plt.figure(figsize=(20,7))
sns.scatterplot(data=df, x="층", y="물건금액")
plt.title("층수와 물건금액의 관계")
plt.show()
df['층'].value_counts().head()
'''
1.0    92724
2.0    80813
3.0    76032
4.0    65998
5.0    51319
'''
```

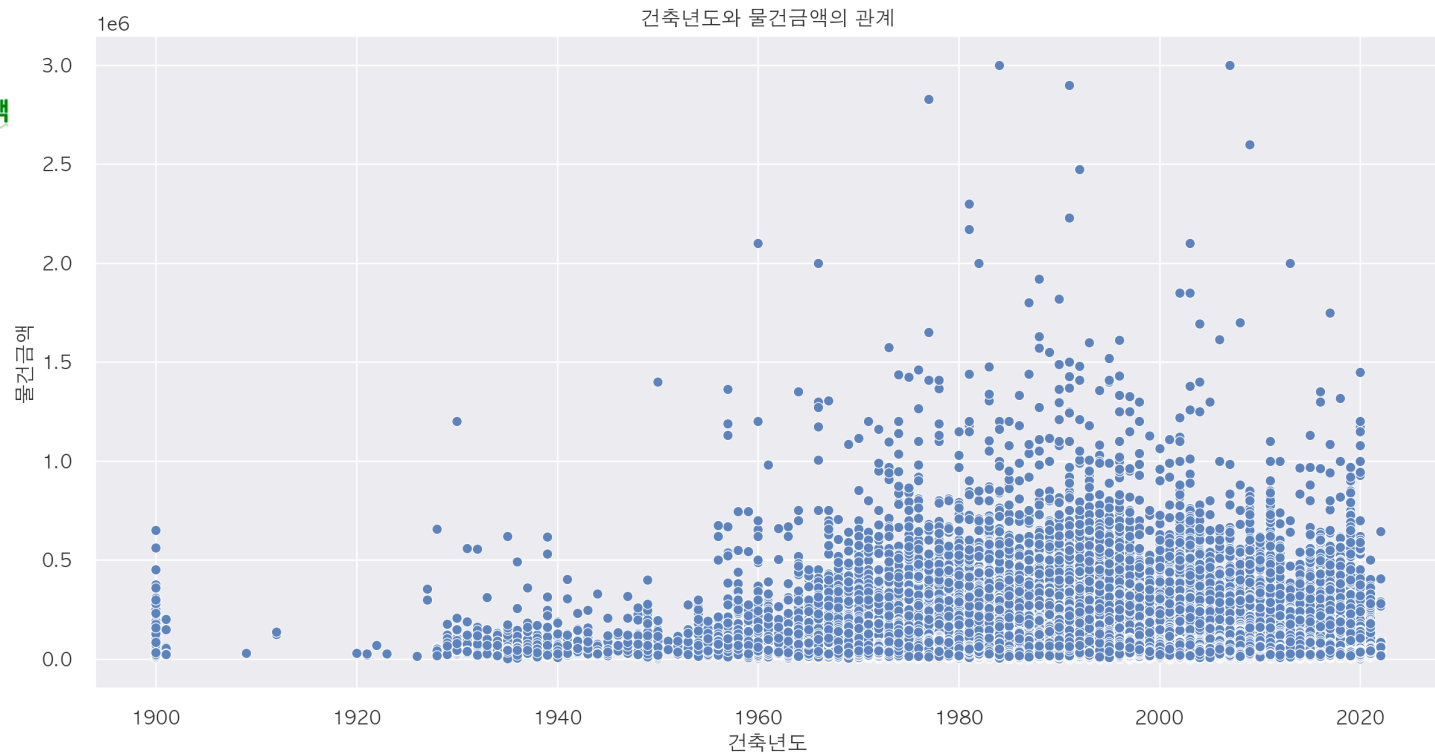


층수에 따라 부동산 거래가격이 변화 없음, 가격이 높은 층은 대부분 1층(단독주택)임

# 3 데이터 시각화 및 분석 : 건축년도 vs 물건금액

# 7) 건축년도 vs 물건금액

```
plt.figure(figsize=(20, 7))  
sns.scatterplot(data=df, x="건축년도", y="물건금액")  
plt.title('건축년도와 물건금액의 관계')  
plt.show()
```



- 건축년도에 따른 부동산 가격에 대한 지표를 보는 것은 의미가 없음
- 1950년대 한국전쟁, 1960년대 아파트 공급 시작을 비롯해 입지별로 가격차이가 나타남
- 신축 건물이라 해서 가격이 비싼 것은 아님



# 3 데이터 시각화 및 분석 : 연속형 변수 기준 상관관계수 확인

# 연속형 변수를 대상으로 상관관계수 확인 및 시각화

```
corr = df.corr(method = 'pearson')
```

```
'''
```

```
      접수연도  물건금액  건물면적  층  건축년도
접수연도  1.000000 -0.011696 -0.056327 -0.056069  0.140506
물건금액 -0.011696  1.000000  0.645417  0.168094 -0.179911
건물면적 -0.056327  0.645417  1.000000 -0.046254 -0.254733
층        -0.056069  0.168094 -0.046254  1.000000  0.168133
건축년도  0.140506 -0.179911 -0.254733  0.168133  1.000000
'''
```

```
heat = sns.heatmap(corr, cbar=True, annot=True,
                    annot_kws={'size':20}, fmt='.2f',
                    square=True, cmap='Blues')
```



상관계수로 보았을 때 물건금액과 건물면적의 상관성이 가장 높음

