

실습과제 1

빅데이터 6기

이신영

요약

본 분석 보고서에서는 한국자산관리공사의 국유부동산 매각현황 데이터를 통해 연도별, 지역별로 매각 특성을 확인하고 금리와 부동산 매각 건수 간의 관계를 분석하였다.

처분 구분 별 비율 확인 결과 매각수의 매각이 약 88%의 높은 비율을 차지했고 매각 금액이 가장 많은 팀과 가장 적은 팀의 차이는 '국유재산3팀'과 '해안면무주지국유화지원추진단'의 매각 금액 차이는 약 1조 6천 억 정도로 상당히 크게 나타났다. 연도별 매각금액은 2021년, 2022년이 가장 높았고, 지역별 매각 면적은 강원도, 경기도, 전라남도 등이 상위 5개 지역으로 나타났다. 또한, 금리가 낮은 년도에는 부동산 매각 건수가 증가할 것이라는 가설을 세우고 분석한 결과, 두 요소 간에 강한 음의 상관 관계를 도출하였다.

이 보고서 결과는 국유부동산 매각현황을 바탕으로 부동산 시장의 특성과 금리의 영향을 이해하는 데 도움을 주고 이후 부동산 정책 모델링과 평가에 영향을 줄 것으로 기대한다.

주요 용어

1. 매각면적: 이 보고서에서의 매각면적은 대장면적(제곱미터)이며, 부동산이나 토지 등이 매각되는 과정에서 해당 부동산의 면적을 의미한다.
2. 매각금액: 매각 금액은 부동산이나 토지 등이 매각되는 과정에서 거래된 금액을 나타낸다.

목차

I 서론

1. 연구 주제 및 목적
2. 연구 질문 및 연구 가설
 - 가. 연구 질문
 - 나. 연구 가설

II 연구 방법

1. 데이터 수집 방법
2. 조사 방법 및 사용된 통계적 방법

III 결과

IV 토의

1. 결과 요약
2. 한계점
3. 제언

I 서론

3. 연구 주제 및 목적

본 연구의 주제는 '한국자산관리공사의 국유부동산 매각현황 데이터를 활용한 시계열 및 지역별 특성 분석'이다. 본 연구는 한국자산관리공사의 국유부동산 매각현황 데이터를 기반으로 매각 금액 비교, 연도별 매각 금액 추이 확인, 상위 5개 지역별 매각 면적 확인, 계약 일자별 매각 특징 분석을 진행하였다. 이를 통해 부동산 시장의 특성 및 경제 환경과의 상호작용을 파악하고, 조직의 지속적인 발전을 지원하는 데 목적을 둔다.

4. 연구 질문 및 연구 가설

다. 연구 질문

[필수 분석]
1) 매각 금액이 가장 많은 팀과 하위 팀의 차이 확인하기
2) 연도별 매각금액 추이 확인하기
3) 상위 5개 지역별 매각 면적 확인하기
4) 계약 일자 정보 기준으로, 매각이 많은 때의 특징 알아보기
[선택 분석]
1) 처분 구분별 비율 확인하기
2) 필수-3 항목을 지역 규모에 따라 보정하기
3) 필수-2 항목을 증감액 그래프로 표현하기*

라. 연구 가설

본 연구의 가설은 [필수-4]에 대해서만 작성하였다.

- 1) 가설: 금리가 낮은 년도에는 부동산 매각 건수가 증가할 것이다.

가) 귀무가설 (H_0): 금리와 부동산 매각 건수 간에는 유의미한 상관 관계가 없다.

나) 대립가설 (H_1): 금리가 낮은 년도에는 부동산 매각 건수가 증가한다.

II 연구 방법

본 연구는 [필수-4]에 대해서만 가설을 설정하였으므로 사용된 통계적 방법 또한 [필수-4]에 대해서만 서술하였다.

3. 데이터 수집 방법

본 연구는 공공데이터포털의 '한국자산관리공사_국유부동산 매각현황_20211231.csv'를 수집하여 활용하였다. 데이터 정보는 다음과 같다.

- 연도구분: YYYY 형식의 연도
- 지역구분: 시/도 정보
- 소재지명: 지번
- 재산구분: 건물 또는 토지로 분류
- 공부지목: 용지 분류
- 대장면적(제곱미터): 0 ~ 20만 가량의 범위
- 대장금액(원): 0 ~ 220억 가량의 범위
- 처분구분명: 처분 분류
- 계약일자: YYYY년 형식의 연도
- 부점명, 팀명

4. 조사 방법 및 사용된 통계적 방법

가. [선택-1] 처분 구분별 비율 확인하기

1) 연구 절차

- 가) '처분구분명'을 기준으로 빈도 계산
- 나) 전체 데이터 수 및 비율 계산
- 다) y축을 로그 스케일로 변경하여 시각화

나. [필수-1] 매각 금액이 가장 많은 팀과 하위 팀의 차이 확인하기

1) 연구 절차

- 가) '팀명'을 기준으로 그룹화하여 팀별 매각 금액의 총합 계산
- 나) 매각 금액의 총합이 가장 많은 팀과 가장 적은 팀 식별
- 다) 매각 금액의 총합이 가장 많은 팀과 가장 적은 팀 차이 계산
- 라) 결과를 데이터프레임으로 변경하여 출력

다. [필수-2] 연도별 매각금액 추이 확인하기

1) 연구 절차

가) '연도구분'을 기준으로 연도별 매각금액 합계 계산

나) 결과를 데이터프레임으로 변경하여 출력

라. [선택-3] [필수-2] 항목을 증감액 그래프로 표현하기

1) 연구 절차

가) Seaborn 패키지를 활용하여 [필수-2]의 결과를 선 그래프로 출력

나) y축의 지수 표기법을 해제하고 y축의 눈금을 '천 만원' 단위로 표시

다) 선 그래프 출력

마. [필수-3] 상위 5개 지역별 매각 면적 확인하기

1) 연구 절차

가) '지역구분'을 기준으로 그룹화하여 매각 면적의 합 계산

나) 매각 면적의 합이 많은 상위 5개 지역 식별

다) 결과를 데이터프레임으로 변경하여 출력

바. [선택-2] [필수-3] 항목을 지역 규모에 따라 보정하기

1) 연구 절차

가) '지역구분'을 기준으로 그룹화하여 대장면적의 합과 총 매각 건수 계산

나) 보정된 지역 규모로 평균 대장면적 계산

다) 평균 대장면적의 합이 많은 상위 5개 지역 식별

라) 결과를 데이터프레임으로 변경하여 출력

사. [필수-4] 계약 일자 정보 기준으로, 매각이 많은 때의 특징 알아보기

1) 연구 절차

가) 계약일자를 날짜 형식으로 변환

나) '년도' 열을 추가하여 계약일자의 년도 추출

다) 년도별 총 매각 건수 계산

라) 총 매각 건수가 많은 순으로 정렬

마) 결과를 데이터프레임으로 변경하여 출력

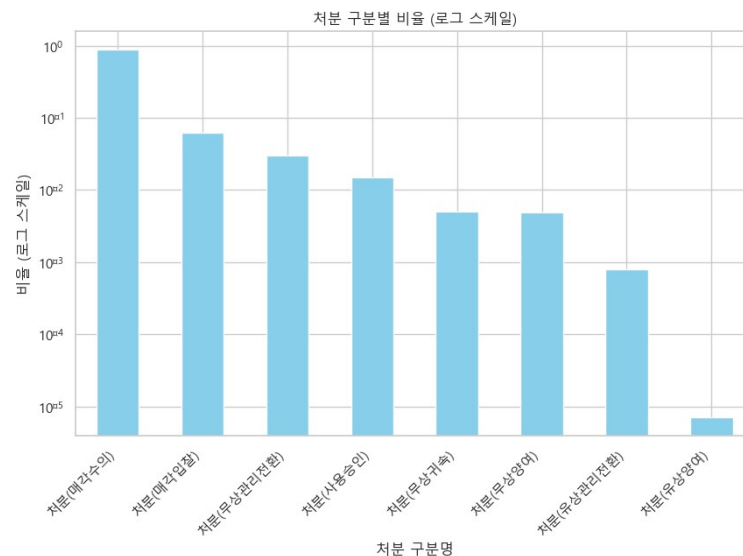
2) 사용된 통계적 방법

금리와 부동산 매각 건수의 상관관계를 파악하기 위해 상관분석을 실시하였다. 년도별 금리 값을 입력한 뒤 총 매각 건수와 금리 간의 상관 계수를 계산하여 이들 간 관계를 확인하고 p-value를 얻어 가설을 검정하였다.

III 결과

1. [선택-1] 처분 구분별 비율 확인하기

처분 구분에 따른 비율 분석 결과, 매각수의 처분이 전체 처분에서 약 88%로 가장 큰 비중을 차지했다. 매각입찰 처분은 약 6%로, 무상관리전환 처분은 약 3%로 상대적으로 낮은 비율을 나타냈다. 매각수의 처분이 차지하는 비율이 압도적으로 크기 때문에 y축을 로그 스케일로 변경하여 막대 그래프로 시각화 하였다.



[그림 1. 처분 구분별 비율]

2. [필수-1] 매각 금액이 가장 많은 팀과 하위 팀의 차이 확인하기

매각 금액의 총합이 가장 많은 팀은 '국유재산3팀(1,581,917,203,990원)', 매각 금액의 총합이 가장 적은 팀은 '해안면무주지국유화지원추진단(22,101,960원)'으로 매각 금액의 총합이 가장 많은 팀과 가장 적은 팀의 차이는 '1,581,895,102,030원'이다

	팀명	대장금액 총합(원)
0	국유재산3팀	1,581,917,203,990
1	국유재산1팀	899,128,271,160
2	국유재산관리2팀	850,425,540,110
3	국유재산2팀	785,818,592,477
4	국유재산관리1팀	733,823,449,808
...
61	원주지부	5,292,361,984
62	춘천지부	4,954,916,510
63	국유통영지부	2,835,397,438
64	기획관리팀	448,202,500
65	해안면무주지국유화지원추진단	22,101,960

[표 2. 팀명별 대장금액 총합]

	팀명	대장금액(원)
	국유재산3팀	1,581,917,203,990원
	해안면무주지국유화지원추진단	22,101,960원
	차이	1,581,895,102,030원

[표 2. 매각 금액이 가장 많은 팀과 가장 적은 팀의 차이]

3. [필수-2] 연도별 매각금액 추이 확인하기

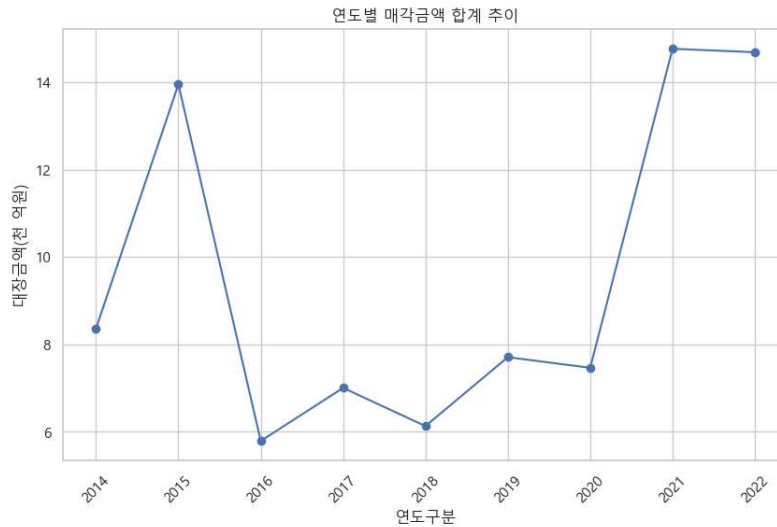
연도별 매각금액 합계 분석 결과는 다음과 같다. 매각금액 합계가 큰 상위 5개년도는 2021년, 2022년, 2015년, 2019년, 2020년이었다.

연도구분	매각금액 합계
2014	836,802,797,181원
2015	1,395,759,106,170원
2016	579,083,546,456원
2017	700,658,462,323원
2018	613,223,822,904원
2019	771,039,378,713원
2020	746,681,020,674원
2021	1,477,228,187,742원
2022	1,469,359,729,072원

[표 3. 연도별 매각금액 합계]

4. [선택-3] [필수-2] 항목을 증감액 그래프로 표현하기

연도별 매각금액 추이를 선그래프로 나타냈다. 2014년에서 2015년에 매각금액이 약 5,600억 증가하고, 2015년에서 2016년에 약 8,200억 급감한 뒤, 2016년에서 2020년에는 약 1,700억 정도 증가했다. 2020년에서 2021년에는 약 7,300억 급증하였고, 2021년에서 2022년에는 약 79억 감소하였다.



[그림 2. 연도별 매각 금액 합계 추이]

5. [필수-3] 상위 5개 지역별 매각 면적 확인하기

매각 면적 총합이 큰 상위 5개 지역은 강원도, 경기도, 전라남도, 충청남도, 경상북도였다.

지역구분	대장면적 합계(제곱미터)
강원도	13355311.950000
경기도	11226483.480000
전라남도	9388504.737000
충청남도	8580609.586000
경상북도	7117240.601000

[표 4. 지역별 대장면적 합계]

6. [선택-2] [필수-3] 항목을 지역 규모에 따라 보정하기

보정된 지역 규모를 고려하여 각 지역별 대장면적을 평균으로 나누어 상대적인 대장면

적을 계산하였다. 그 결과, 평균대장면적의 합계는 강원도, 충청남도, 충청북도, 전라남도, 제주특별자치도 순으로 많았다.

지역구분 평균대장면적 합계	
강원도	1107.313817
충청남도	990.946944
충청북도	764.096355
전라남도	715.423664
제주특별자치도	707.290613

[표 5. 지역별 평균대장면적 합계]

7. [필수-4] 계약 일자 정보 기준으로, 매각이 많은 때의 특징 알아보기

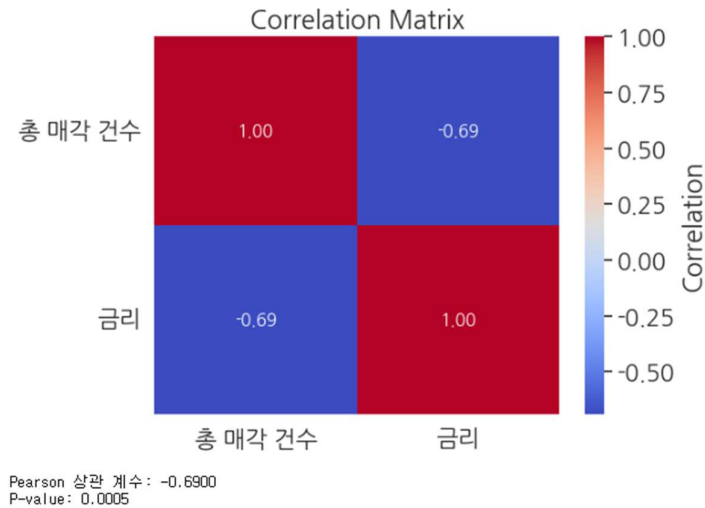
계약 일자 정보는 YYYY년 형식의 연도이므로 연도별 총 매각 건수를 분석한 결과, 매각이 많은 상위 5개년도는 2017년, 2016년, 2015년, 2014년, 2019년이었다.

연도	총 매각 건수
2017	23306
2016	18853
2015	17228
2014	17155
2019	14700
2018	14178
2022	14128
2020	12331
2021	9235
2010	65
2013	42
2011	30
2007	17
2009	15
2008	11
2012	9
2006	4
2000	4
2001	2
2005	2
2002	2

[표 6. 연도별 총 매각 건수]

[필수-4]의 경우, 앞서 설정한 가설 ‘금리가 낮은 년도에는 부동산 매각 건수가 증가할 것이다.’의 검정을 위해 상관 분석을 진행하였다. 상관행렬을 히트맵으로 시각화 한 결과, 총 매각 건수와 금리 간에 부정적인 상관 관계가 나타났다. 상관 계수는 -0.69로 나타났으며, 이는 두 변수 간에 강한 음의 선형 관계가 있음을 시사한다. 분석 결과, 금리가 낮

은 년도에 부동산 총 매각 건수가 상대적으로 높았다는 통계적인 근거를 확인할 수 있었다. 이러한 결과는 부동산 시장에서 금리가 부동산 거래에 미치는 영향을 이해하는 데 도움을 줄 것으로 예상된다. 또한, P-value는 0.0005로, 0.05보다 작으므로 '귀무가설 (H0): 금리와 부동산 매각 건수 간에는 유의미한 상관 관계가 없다.'를 기각할 수 있다.



[그림 3. 금리와 총 매각 건수의 상관분석 결과]

IV 토의

1. 결과 요약

이 보고서에서는 한국자산관리공사의 국유부동산 매각현황 데이터를 활용하여 분석을 진행하였다. 본 분석 결과는 다음과 같이 요약할 수 있다.

첫째, [선택-1]을 통해 처분 구분별 비율을 확인한 결과, 매각수의 처분이 압도적으로 큰 비율을 차지하였다. 둘째, [필수-1]을 통해 매각 금액이 가장 많은 팀인 '국유재산3팀'과 가장 적은 팀인 '해안면무주지국유화지원추진단'의 차이를 확인한 결과, 약 1조 6천억의 큰 차이를 보였다. 셋째, [필수-2]를 통해 연도별 매각금액 추이를 확인한 결과, 매각금액 합계가 큰 연도는 2021년, 2022년 등이었다. 넷째, [선택-3]을 통해 [필수-2] 항목을 증감액 그래프로 나타냄으로써 연도별 매각금액 추이를 시각화 하여 금액의 증감 패턴을 확인할 수 있었다. 다섯째, [필수-3]을 통해 상위 5개 지역별 매각 면적을 확인한 결과, 상위 5개 지역은 강원도, 경기도 등으로 나타났다. 여섯째, [선택-2]를 통해 [필수-3] 항목을 지역 규모에 따라 보정한 결과, 상위 지역은 강원도, 충청남도 등이었다. 일곱째, [필수-4]를 통해 계약 일자 정보를 기준으로 매각이 많은 때는 2017년, 2016년 등이었다. 이 결과를 바탕으로 금리와 부동산 매각 건수의 상관 분석을 진행하였고, 두 변수 간에 강한 음의 상관 관계가 있음을 확인함으로써 통계적으로 유의미한 결과를 얻었다.

2. 한계점

본 분석의 한계는 다음과 같다.

첫째, 상관 분석은 인과관계를 직접적으로 나타내지 않으며, 다양한 외부 요인들이 결과에 영향을 미칠 수 있으므로 통계적 한계를 갖는다. 둘째, 부동산 시장은 복잡하고 다양한 영향 요인들이 존재하므로, 이를 모두 고려하기 어렵다는 데이터 한계를 갖는다. 셋째, 본 분석은 특정 지역이나 시기에 대한 분석이므로 다른 지역이나 시기에서는 결과가 달라질 수 있기 때문에 분석의 일반성에서의 한계를 갖는다.

3. 제언

이 보고서에서는 공공데이터, 그 중에서도 부동산 데이터를 수집하여 다각적으로 분석하였다. 추후 가능한 분석 방향은 다음과 같다.

먼저, 부동산 시장의 특수성을 고려한 모델링으로 추가적인 인자들을 고려한 심층 분석이 가능하다. 그리고 부동산 거래와 관련된 정책의 효과를 평가하여 향후 정책 제안에 활용할 수 있도록 정책적 시사점을 도출할 수 있다. 또한, 시계열 데이터를 이용하여 시간 경과에 따른 변동을 더욱 정확히 파악하는 것이 중요하다. 마지막으로 지역별 특성을 고려한 추가 분석을 통해 지역 간 부동산 시장의 특이성을 파악할 필요가 있다.

V 참고문헌

전해정. "부동산 뉴스와 아파트 매매가격과 거래량 간의 관계에 대한 빅데이터 시계열 분석." 不動産法學 24.2 (2020): 53-69.

강중용,동재욱,and 이화룡. "공공기관 종전부동산의 매각과 부동산거래량의 상관분석-서울지역을 중심으로-." 大韓不動產學會誌 35.1 (2017): 127-145.

한국은행 경제통계시스템[웹사이트].(2023.12.21).URL: <https://ecos.bok.or.kr/#/SearchStat>