

# 부동산 전세가격 예측·전세가율 분석

: 적정 전세가율을 활용한 전세사기 예방 웹사이트 구축

Statistical Analysis & Machine Learning

## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)

주제 선정 배경

프로젝트 수행 방향

프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집

데이터 전처리

EDA

통계 분석

변수 설정

분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가

서울 자치구별 전세가격 예측

Web 구축 & 배포

### 04. 최종 결론

기대 효과

한계점

### 05. Document

데이터 정의서

참고문헌

출처

# 부동산 전세가격 예측·전세가율 분석

## 프로젝트 조직 (구성원 및 역할)



**김주성(팀장)**

데이터 전처리, 통계 분석, 머신러닝



**이길연**

데이터 전처리, EDA, 통계 분석



**최동원**

데이터 전처리, 머신러닝



**최선재**

데이터 전처리, 시각화, 웹연동



**최연우**

데이터 전처리, EDA, 시각화



**황신엽**

데이터 전처리, EDA

# 부동산 전세가격 예측·전세가울 분석

주제 선정 배경

국내

## "전세사기 막을 보완대책 시급"

'전세사기 특별법 보완 및 문제 해결을 위한 정책 포럼' 주제로 국회토론회 개최

23.07.18 17:38 | 최종 업데이트 23.07.19 09:06 | 윤종은(tomayoun) ▼

전세사기 특별법 제정 과정에서 피해자들의 의견이 반영되지 못했다는 응답은 85.4%로 대부분 소외감을 나타냈으며 특별법에 따른 피해구제책에 대해 91%가 충분하지 않다고 응답했다. 또 구제책을 아예 모르고 있거나 정책 내용을 이해하지 못하고 있으며 정책이 도움이 안 된다고 응답한 경우가 많았다.

연합뉴스 구독

## 무자본 갭투자로 126채 산 뒤 전세보증금 250억 편취한 일당

입력 2023.08.02. 오전 10:06 | 수정 2023.08.02. 오전 10:10 | 기사원문

경기북부경찰청은 부동산 업체 대표 A씨 등 11명을 검거, 8명을 구속했다. 이들은 무자력자를 내세워 주택 126채를 허위 매수하고 전셋값과 매매가를 유사하게 설정하여 전세보증금 약 253억원을 편취한 혐의를 받는다. 조직적으로 역할을 분담하여 세입자들에게 신뢰를 줬으며, 일부는 가짜 전세계약서로 청년전세대출도 받았다.

▶ 갭 투자를 비롯하여 여러 전세사기 문제가 발생

→ 부동산 시장에 불신 증가 & 불안정한 전세가격 형성에 기여

# 부동산 전세가격 예측·전세가율 분석

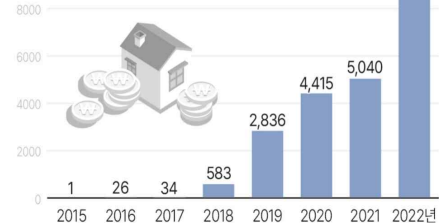
## 주제 선정 배경

국내

### HUG 전세보증금 대위변제액 추이

HUG가 집주인을 대신해 세입자에게 갚아준 전세보증금

단위: 억원



연암뉴스

자료: 주택도시보증공사(HUG)

사회

### 신혼부부 꿈 망친 전세사기... "소송 다 이겼지만 소용없어"

2023년 04월 26일 05시 18분 댓글

### "경매 막겠다는 말 믿었는데" 울산 오피스텔 임차인들, 전세사기 고소

남구 D오피스텔 58세대 강제경매...임대인·회사 대표 고소 예정

(울산=뉴스1) 임수정 기자 | 2023-08-29 15:59 송고

이어 6월 배씨는 "권씨는 상환 능력이 없고 실소유주인 본인이 직접 이자를 갚아 경매를 취하하겠으니 보증금 만기를 6개월 연장해달라"는 취지로 임차인들을 설득했다고 한다. 배씨의 말을 믿고 기다리다 강제경매 소식을 들은 임차인들은 허탈감을 감추지 못했다.

임차인 김모씨는 "임차인들 대부분이 20대, 30대 초반으로 부동산이나 법률에 대해 잘 알지 못하고 전세보증금반환보증에도 가입하지 않았다"며 "배씨가 도의적 책임을 지고 사태를 해결하겠다는 말을 믿으면서 아무 대처도 하지 않고 흘러보낸 시간이 후회된다"고 말했다.

H 부동산임대회사 대표 배씨는 수 차례 연락을 시도했으나 연락이 닿지 않았다.

(1) 전세사기 피해 + 전세보증금 대위변제액 급증

(2) 전세사기 피해자가 승소하더라도 임대인이 '배째라는 식'의 행태로 대응하면 마땅한 해결책이 없는 상황

→ 우선적으로 전세사기 예방이 가장 중요하다

# 부동산 전세가격 예측·전세가율 분석

주제 선정 배경

해외

## 파월은 물론 연준 간부들 잇달아 추가 금리인상 시사

뉴스1 원문 | 기사전송 2023-08-28 06:47 최종수정 2023-08-28 08:48

파월 "물가 여전히 높아" : 파월 의장은 지난 25일 잭슨홀 회의 개막 연설에서 "소비자 지출이 견조하고, 주택 부문이 반등할 가능성이 있는 가운데 경제가 예상만큼 냉각되지 않을 수 있다는 징후에 주의를 기울이고 있다"고 말했다. 이에 따라 "적절하다고 판단되면 금리를 더 올릴 준비가 돼 있으며, 물가 상승률이 목표치를 향해 내려가고 있다는 확신이 들 때까지 긴축 정책을 계속 유지할 생각"이라고 덧붙였다.

MTO머니투데이 (+) 구독

PICK 1

## 中 부동산위기 진앙지 헝다, 상반기만 6조 손실..."주식 거래재개"

입력 2023.08.28. 오전 9:30 · 수정 2023.08.28. 오전 9:31 기사원문

중국의 대형 부동산 개발업체인 비구이위안(碧桂園·컨트리가든)이 디폴트 위기를 맞으면서 중국 부동산 업계에 '도미노 디폴트' 우려가 번지고 있다. 다른 중국 주요 부동산 업체의 재정 상황도 나쁜 상황이다. 중국 부동산 시장은 중국 국내총생산(GDP)의 25%를 차지한다. 부동산 시장이 흔들리면 중국 경제도 위기에 빠질 수 있다는 의미다.

미국 금리인상 지속 + 중국 부동산 리스크 등

→ 세계 경제 정세의 불확실성 또한 국내 부동산 시장의 불안정에 영향을 끼침

## 주제 선정 배경

### 주식회사설립 등기신청서

- 상 호 : 건물주 주식회사
  - 본 점 : ○○ ○○구 ○○동 ○○○번지
  - 등기의 목적 : 주식회사 설립등기
  - 등기의 사유 : 정관을 작성하고 공증인의 인증을 받아 발기인이 회사 설립시에 발행하는 주식의 총수를 인수받고, 2023년 9월 26일 상법 제299조의 절차를 종료하였으므로 다음 사항의 등기를 구함.  
상 호 : 건물주 주식회사  
본 점 : ○○ ○○구 ○○동 ○○번지  
목 적 : 1) 아파트 전세 가격 예측  
2) 정부의 부동산 안정화 정책의 기여  
3) 개인별 맞춤 투자전략 구축  
4) 위 각호에 관련된 부대사업
- 회사가 발행할 주식의 총수 : 40,000주  
1주의 금액 : 금 5,000원

회사의 설립시에 발행하는 주식의 총수 : 10,000주  
발행주식의 총수, 그 종류와 각종 주식의 내용과 수 : 10,000주 보통주식  
자본의 총액 : 금 50,000,000원  
광고방법 : 서울특별시 내에서 발행하는 일간 ○○일보에 게재한다.  
회사의 존립기간 또는 해산사유 : 회사성립일로부터 만 30년  
이사의 성명 이사 이 길 연  
이사 최 동 원  
이사 최 선 재  
이사 최 연 우  
이사 황 신 업

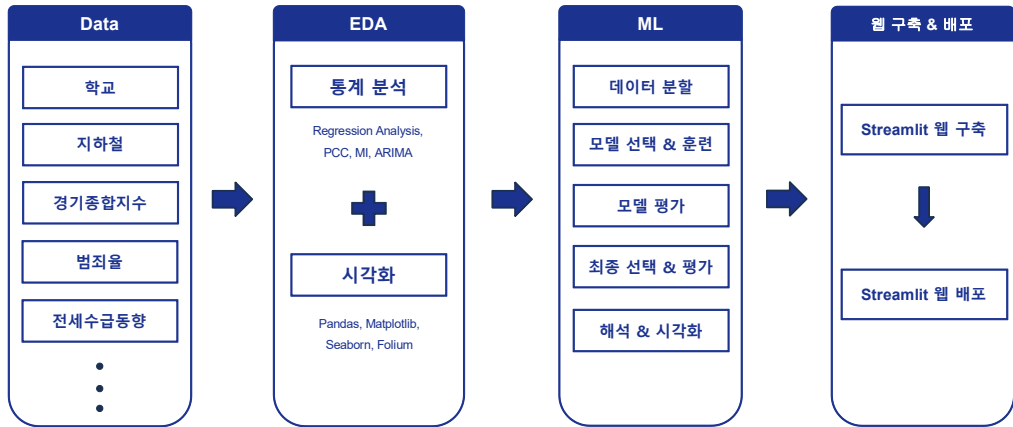
회사를 대표할 이사의 성명과 주소  
대표이사 김 주 성

위와 같이 등기신청합니다

2023년 9월 26일  
신청인 건물주 주식회사  
김 주 성 (인)  
(이하 이사 전원이 기명날인한다)

# 부동산 전세가격 예측·전세가율 분석

## 프로젝트 수행 방향





### Frameworks

#### Statistics



PCC



PROPHET

#### Machine Learning



Random Forest

XGBoost



LightGBM

#### Deep Learning



TensorFlow



Keras

ANN

LSTM

## 프로젝트 수행 방향

### Statistical Analysis

Regression Analysis, PCC, Prophet

(OLS, R-Squared, Correlation, Prediction)

### ML

Random Forests, LightGBM, XGBoost

(Ensemble Algorithms)

### DL

ANN, LSTM

(Neural Network Architectures)



부동산 전세가격  
예측 · 분석

## 프로젝트 수행 방향 - 변수 설정

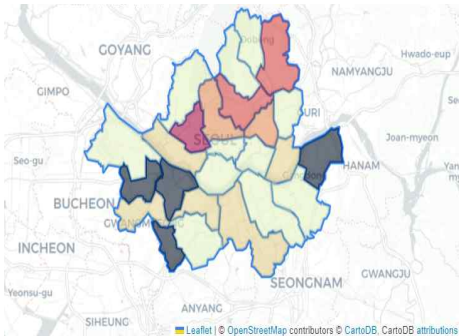


**데이터 전처리**

# 부동산 전세가격 예측·전세가율 분석

## 프로젝트 수행 방향 - 평가 지표

### 지도 시각화



### 공간정보 분석

#### 데이터 시각화

자치구별 위치 데이터

공원 위치 데이터

학교 위치 데이터

지하철역 위치 데이터



#### 구역 설정

서울 자치구별로 분할 후 데이터 분석

#### 기대 효과

1. 서울 자치구별로 적정 전세가격 확인
2. 전세사기 의심 매물 추정 가능

# 부동산 전세가격 예측·전세가율 분석

## 프로젝트 추진 일정

구분	1주차(9/6 ~ 9/10)			2주차(9/11 ~ 9/17)			3주차(9/18 ~ 9/24)			4주차(9/25 ~ 9/26)	
분석 기획											
데이터 수집											
논문 자료 수집											
데이터 전처리											
탐색적 데이터 분석											
데이터 시각화											
통계 분석											
모델 생성 및 평가											
Streamlit 배포											
웹 연동 및 배포											
프로젝트 발표											

## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)

주제 선정 배경

프로젝트 수행 방향

프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집

데이터 전처리

EDA

통계 분석

변수 설정

분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가

서울 자치구별 전세가격 예측

Web 구축 & 배포

### 04. 최종 결론

기대 효과

한계점

### 05. Document

데이터 정의서

참고문헌

출처

# 부동산 전세가격 예측·전세가율 분석

## 데이터 수집

기존 연구를 참고하여 경제 지표와 주거 특성이 주택 전세가격에 영향을 준다는 것을 파악

따라서 기준 금리, 실업률, 전세수급동향, 주택 매매가격지수, 경기종합지수, 지하철역 위치, 학교(초, 중, 고등학교) 위치, 대학 위치 변수 채택

지역주택가격 변동의 장단기 결정요인에 관한 실증분석.

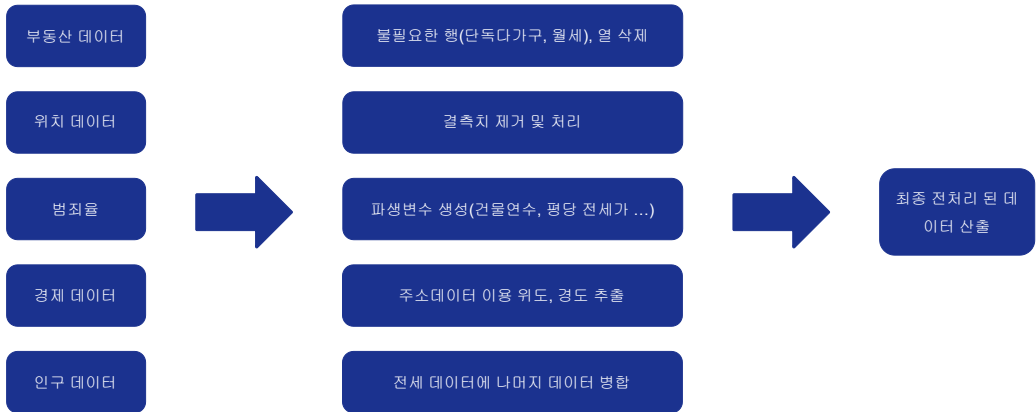
구분	변수명	변수정의 (단위)
종속 변수	주택가격	지역별 주택매매가격종합지수 (2015.12=100)
	금리	한국은행 기준금리 (%)
설명 변수 (전국)	주식가격	KOSPI 종합지수 (월평균, 1980.1.4=100)
	주택대출	예금은행주택대출 (말잔, 십억원)
	경기상황	경기동행지수 (2010=100)
	물가	지역별 소비자물가지수 (2010=100)
설명 변수 (지역)	주택보증	지역별 주택금융신용보증 공급 (말잔, 억원)
	전세가격	지역별 주택전세가격종합지수 (2015.12=100)
	취업자수	지역별 총취업자수 (천명)
	가구수	지역별 주민등록세대수 (세대)
	고령인구 비중	지역별 65세 이상 인구 비중 (%)

주택시장 경기변동과 주거특성들의 아파트가격에 대한 영향력 변화  
- 용인 동백택지개발지구를 사례 -

구 분		요 인	비 고
주거 특성	개별적 특성	대지지분	규모를 측정하기 위한 변수로서 다중공선성 관계에 있기 때문에 둘 중 하나만을 이용
		평형	
		층	주택의 해당 층
		각도(향)	정남향에 근접한 정도
	단지적 특성	평균층수	단지 내 주거동의 평균층수
		세대수	단지 내 총 주택수
		브랜드 더미	LH=0 민영=1
	접근성	중심업무지구까지의 거리	동백지구에서는 다중공선성 관계에 있기 때문에 둘 중 하나만을 이용
		중앙공원까지의 거리	
		경전철역까지의 거리	-
		초등학교까지의 거리	-
		고속도로IC까지의 거리	-
복합행정타운까지의 거리		-	

이외에도 주택 전세가격에 미치는 다른 요인들을 확인하기 위해 범죄율(치안), 인구수 변수를 채택

## 데이터 전처리





# 부동산 전세가격 예측·전세가율 분석

## 데이터 전처리

Before

	접수 년도	자치구 코드	자치구 명	법정동 코드	법정동 명	지번구 분코드	지번구 분	본번	부번	층	...	보증금 (만원)	임대 료(만원)	건물명	건축년 도	건물 종도	계약기간	신규계 약구분	갱신청 구권사 유	중전보 증금	중전임 대료
5452627	2023	11710	송파구	111000	방이동	1.0	대지	28.0	2.0	5.0	...	5000	57	엘루이 시티	2019.0	오피스 빌딩	NaN	NaN	NaN	NaN	NaN
5452628	2023	11280	중랑구	101000	명파동	NaN	NaN	NaN	NaN	NaN	...	6000	0	NaN	1992.0	단독 다가구	NaN	NaN	NaN	NaN	NaN
5452629	2023	11710	송파구	104000	송파동	1.0	대지	141.0	20.0	5.0	...	2000	62	해피하 우스	2014.0	연립 다세대	23.01~25.01	신규	NaN	0.0	0.0
5452630	2023	11590	동작구	102000	신도동	NaN	NaN	NaN	NaN	NaN	...	1000	25	NaN	1990.0	단독 다가구	NaN	NaN	NaN	NaN	NaN
5452631	2023	11350	노원구	103000	공릉동	1.0	대지	714.0	0.0	13.0	...	20000	0	삼익4단 지아파트	1994.0	아파트	23.01~26.01	신규	NaN	0.0	NaN



After

	접수 년도	자치구 코드	자치구 명	법정동 코드	법정동 명	지번구 분코드	지번구 분	본번	부번	층	계약일	전월 세구분	임대 면적	보증금 (만원)	임대료 (만원)	건물명	건축년 도	건물 종도
2439177	2023	11350	노원구	105000	상계동	1.0	대지	624.0	0.0	6.0	20221108	전세	49.94	15280	0	상계주공 15(고층,공무 원임대)	1988.0	아파트
2439178	2023	11710	송파구	102000	신천동	1.0	대지	17.0	0.0	20.0	20221107	전세	84.79	88000	0	팍크리오	2008.0	아파트
2439179	2023	11710	송파구	112000	오륜동	1.0	대지	165.0	0.0	4.0	20221106	전세	45.77	30000	0	삼익2차아파트	1988.0	아파트
2439180	2023	11740	강동구	103000	상일동	1.0	대지	515.0	0.0	19.0	20221106	전세	84.85	65000	0	고막숲아이파크	2018.0	아파트
2439181	2023	11350	노원구	103000	마사동	1.0	대지	714.0	0.0	13.0	20221105	전세	39.60	20000	0	삼익4단지아파트	1994.0	아파트

### 불필요한 열 삭제

- 전세 가격 예측에 불필요한 열 삭제
- '계약기간, 신규계약구분, 갱신청구권사용, 중전보증금, 중전임대료' 컬럼 삭제

```
columns_to_drop = ['계약기간', '신규계약구분', '갱신청구권사용', '중전보증금', '중전임대료']
seoul_js_data = seoul_js_data.drop(columns = columns_to_drop)
```

### 불필요한 행 삭제

- '건물용도 = 단독다가구'인 행 삭제 (본번, 부번이 없어 위도, 경도 매칭이 불가)
- '전월세구분 = 월세'인 행 삭제 (전세 가격 예측 프로젝트임으로)

```
seoul_js_data = seoul_js_data[seoul_js_data.건물용도 != '단독다가구']
seoul_js_data = seoul_js_data[seoul_js_data.전월세구분 != '월세']
```

# 부동산 전세가격 예측·전세가율 분석

## 데이터 전처리

Before

접수 년도	자치구 코드	자치구 명	법정동 코드	법정동 명	지번 구분 코드	지번 구분 명	본번	부번	층	계약일	전월 세구 분	임대 면적 (㎡)	보증금 (만원)	임대료 (만원)	건물명	건축년 도	건물 종류
2439177	2023	11350	노원구	상계동	1.0	대지	624.0	0.0	6.0	20221108	전세	49.94	15280	0	상계주공 15고층,공무원임대	1988.0	아파트
2439178	2023	11710	송파구	신천동	1.0	대지	17.0	0.0	20.0	20221107	전세	84.79	88000	0	파크리오	2008.0	아파트
2439179	2023	11710	송파구	오금동	1.0	대지	165.0	0.0	4.0	20221106	전세	45.77	30000	0	상아2차아파트	1988.0	아파트
2439180	2023	11740	강남구	상일동	1.0	대지	515.0	0.0	19.0	20221106	전세	84.85	65000	0	고덕솔아이파크	2018.0	아파트
2439181	2023	11350	노원구	민재동	1.0	대지	714.0	0.0	13.0	20221105	전세	39.60	20000	0	삼익4단지아파트	1994.0	아파트



After

접수 년도	자치구 코드	자치구 명	법정동 코드	법정동 명	지번 구분 코드	지번 구분 명	본번	부번	층	...	임대 면적 (㎡)	보증금 (만원)	임대 료 (만원)	건물명	건축년 도	계약 년도	계약 월	건물 종류	주소
2439177	2023	11350	노원구	상계동	1.0	대지	624	0	6.0	...	49.94	15280	0	상계주공 15고층,공무원임대	1988.0	2022	11	아파트	노원구 상계동 624
2439178	2023	11710	송파구	신천동	1.0	대지	17	0	20.0	...	84.79	88000	0	파크리오	2008.0	2022	11	아파트	송파구 신천동 17
2439179	2023	11710	송파구	오금동	1.0	대지	165	0	4.0	...	45.77	30000	0	상아2차 아파트	1988.0	2022	11	아파트	송파구 오금동 165
2439180	2023	11740	강남구	상일동	1.0	대지	515	0	19.0	...	84.85	65000	0	고덕솔아이파크	2018.0	2022	11	아파트	강남구 상일동 515
2439181	2023	11350	노원구	민재동	1.0	대지	714	0	13.0	...	39.60	20000	0	삼익4단지아파트	1994.0	2022	11	아파트	노원구 민재동 714

## 파생변수 생성

- '계약일' 컬럼으로 '계약년도, 계약월' 컬럼 생성
- '계약년도 - 건축년도'로 '건물연수' 컬럼 생성
- '발생합계 / 인구수'로 '범죄율' 컬럼 생성
- '자치구명 + 법정동명 + 본번 + 부번'으로 '주소' 컬럼 생성

```
seoul_js_data['계약일'] = pd.to_datetime(seoul_js_data['계약일'], format='%Y%m%d')
seoul_js_data['계약년도'] = seoul_js_data['계약일'].dt.year
seoul_js_data['계약월'] = seoul_js_data['계약일'].dt.month
seoul_js_data['건물연수'] = seoul_js_data['계약년도'] - seoul_js_data['건축년도']
Crime_df['범죄율'] = Crime_df['발생합계'] / Crime_df['인구수'] * 100
```



# 부동산 전세가격 예측·전세가율 분석

## 데이터 전처리

Before

	연도	자치구 코드	자치구 명	면적(㎡)	법정동명	지번구분 코드	지번구분	면적	부속	총 ...	임대료 (만원)	건물명	건축년도	건물용도	계약년도	계약월	건물연수	주소	위도	경도	
0	2011	11170	동산구	11300.0	원로포구	1.0	대지	1	0	5.0	-	0	(T)	1990.0	오피스텔	2011	4	21.0	동산구 원로포구 1-1	37.537739	126.965287
1	2011	11170	동산구	11300.0	원로포구	1.0	대지	1	0	6.0	-	0	(T)	1990.0	오피스텔	2011	10	21.0	동산구 원로포구 1-1	37.537739	126.965287
2	2011	11440	마포구	11200.0	노고산동	1.0	대지	1	1	10.0	-	0	(I-1)	1998.0	오피스텔	2011	2	13.0	마포구 노고산동 1-1	37.556254	126.942750
3	2011	11440	마포구	11200.0	노고산동	1.0	대지	1	1	5.0	-	0	(I-1)	1998.0	오피스텔	2011	11	13.0	마포구 노고산동 1-1	37.556254	126.942750
4	2011	11650	서초구	10100.0	병내동	1.0	대지	1031	3	3.0	-	0	(1031-3)	1987.0	오피스텔	2011	7	24.0	서초구 병내동 1031-3	37.478172	126.999450

After

	계약 년도	계약 월	자치구 코드	자치구 명	법정동 코드	법정동 명	지번구 분	지번구 명	면 적	부 속	...	선행종합 합치수	동행종합 합치수	후행종합 합치수	전세수급 동향	서울 실 거래가격 치수	법정 동	위도	경도	평균매매 가격	전세 가격
24391177	2022	11	11350	노고산구	10500.0	상지동	1.0	대지	624	0	-	109.7	108.9	111.8	72.039914	146.8	NaN	37.668783	127.053725	NaN	15280
24391178	2022	11	11710	송파구	10200.0	신천동	1.0	대지	17	0	-	109.7	108.9	111.8	72.039914	146.8	NaN	37.520536	127.106151	165000.0	88000
24391179	2022	11	11710	송파구	11200.0	오륜동	1.0	대지	165	0	-	109.7	108.9	111.8	72.039914	146.8	NaN	37.497938	127.136728	NaN	30000
24391180	2022	11	11740	강남구	10300.0	상도동	1.0	대지	515	0	-	109.7	108.9	111.8	72.039914	146.8	NaN	37.551493	127.161300	NaN	65000
24391181	2022	11	11350	노고산구	10300.0	상지동	1.0	대지	714	0	-	109.7	108.9	111.8	72.039914	146.8	NaN	37.629680	127.071928	NaN	20000

## 전세데이터에 나머지 데이터 병합

- 인구 데이터 병합
- 기준금리, 실업률 데이터 병합
- 경기종합지수 데이터 병합
- 전세수급동향 데이터 병합
- 매매 실거래가격지수, 평균매매가격 데이터 병합

```
merged_df = pd.merge(merged_df, population_df, on=['계약년도', '자치구명'], how='left')
merged_df = pd.merge(merged_df, IR_LR_df, on=['계약년도', '계약월'], how='left')
merged_df = pd.merge(merged_df, EI_df, on=['계약년도', '계약월'], how='left')
merged_df = pd.merge(merged_df, NSDT_df, on=['계약년도', '계약월'], how='left')
merged_df = pd.merge(merged_df, HSP1_df, on=['계약년도', '계약월'], how='left')
merged_df = pd.merge(merged_df, seoul_sell_group, on=['계약년도', '계약월', '주소', '건물용도'], how='left')
```

## 데이터 전처리

Before

After



### 결측치 제거

· '건축년도, 층, 본번, 부번, 법정동명, 평균매매가격'의 결측치 제거

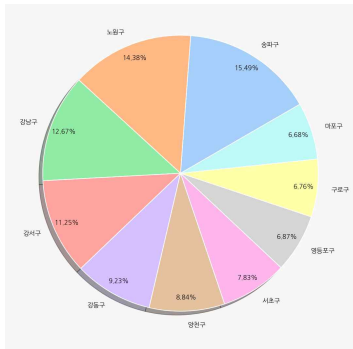
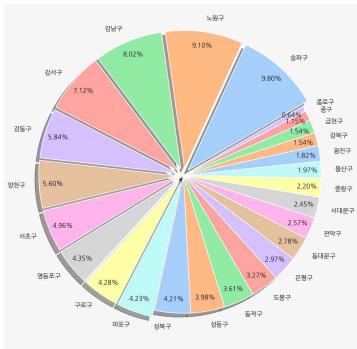
```
seoul_js_data.dropna(subset=['건축년도'], inplace = True)
seoul_js_data.dropna(subset=['층'], inplace = True)
seoul_js_data.dropna(subset=['본번'], inplace = True)
seoul_js_data.dropna(subset=['부번'], inplace = True)
seoul_js_data.dropna(subset=['법정동명'], inplace = True)
```

### 결측치 처리

- '기준금리' 결측치(23년 8,9월)는 '3.5'로 대체
- '범죄율' 결측치는 연도별, 자치구별 평균으로 대체
- XGBoost 모델 활용 결측치를 예측하여 대체

# 부동산 전세가격 예측·전세가율 분석

EDA



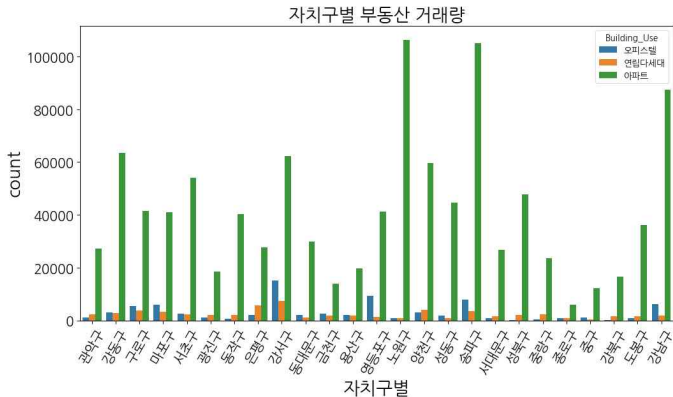
서울 자치구별 부동산 거래 데이터 비율

2011년도부터 2023년 7월까지의 서울 자치구별 부동산 거래량을 파이차트로 나타냄  
 송파구, 노원구, 강남구, 강서구, 강동구, 양천구, 서초구, 영등포구, 구로구, 마포구가 자치구별 거래량 Top10에 해당하는 것으로 발견

서울 자치구별 부동산 거래 데이터 비율 Top 10

서울 거래량 Top10 자치구별 중 송파구, 노원구, 강남구, 강서구가 많은 거래량을 차지하고 있음을 발견

## EDA

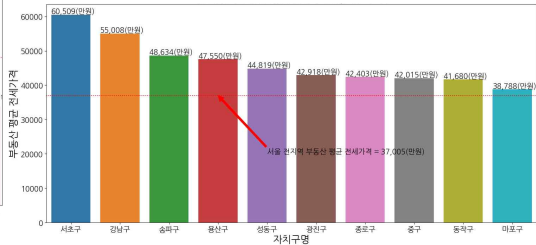
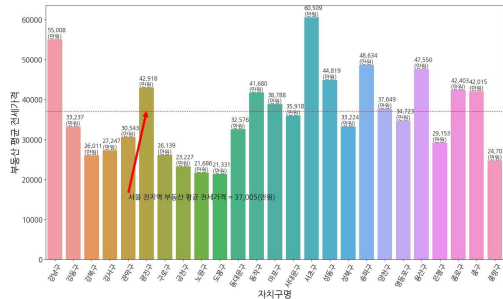


### 전세데이터에 나머지 데이터 병합

- 인구 데이터 병합
- 기준금리, 실업률 데이터 병합
- 경기종합지수 데이터 병합
- 전세수급동향 데이터 병합
- 매매 실거래가격지수, 평균매매가격 데이터 병합

# 부동산 전세가격 예측·전세가율 분석

EDA



서울 자치구별 부동산 평균/중위 전세가격

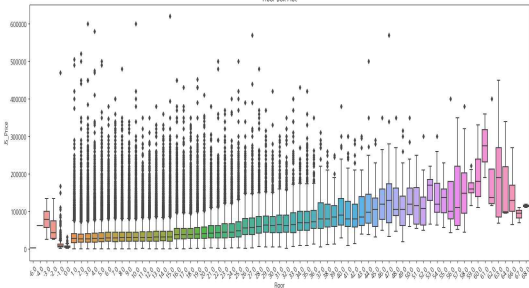
서울 자치구별 부동산 평균/중위 전세가격 Top 10



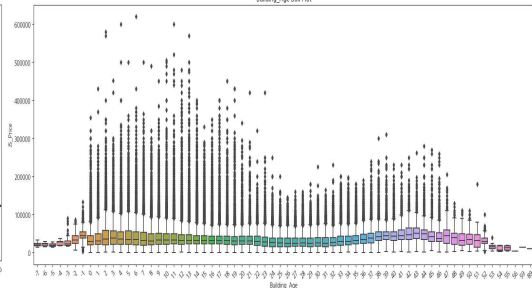
# 부동산 전세가격 예측·전세가율 분석

## EDA

Floor Box Plot



Building\_Age Box Plot



### 층 Box Plot

- 층이 높아질수록 전세가격이 상승하는 것을 확인

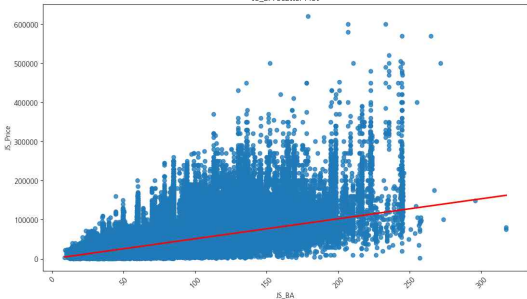
### 건물연수 Box Plot

- 건물연수(계약년도-건축년도)가 적거나 높을 때 전세가격이 높은 것을 확인
- 건물연수가 높을 때 전세가격이 높은 것은 재개발에 대한 기대효과라고 추측

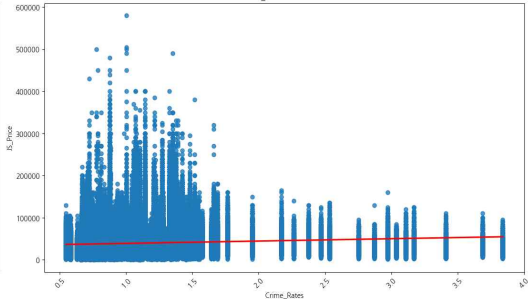
# 부동산 전세가격 예측·전세가율 분석

## EDA

JS\_BA Scatter Plot



Crime\_Rates Scatter Plot



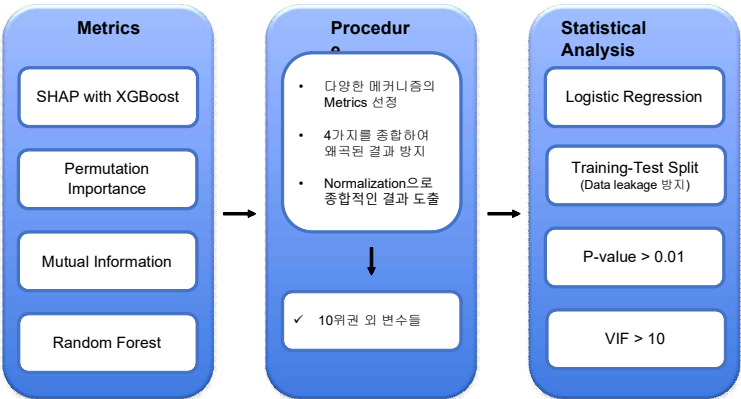
### 임대면적 산점도

임대면적이 넓을수록 전세가격이 상승하는 것을 확인

### 범죄율 산점도

일반적 통념과 달리 범죄율 높을수록 전세가격이 상승하는 것을 확인

## Feature Ranking



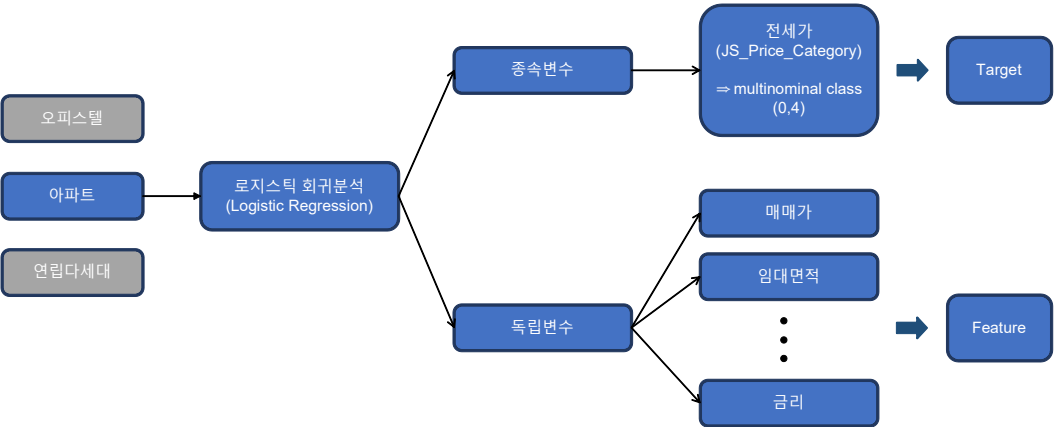
10위권 외 변수들

## Feature Ranking

	Feature	Combined_Ranking
12	Sell_Price	0.986482
7	JS_BA	0.684255
0	Building_Age	0.251346
19	latitude	0.189964
20	longitude	0.181056
10	Region_Name	0.145718
9	Population	0.136831
15	Shortest_Distance_to_University	0.124809
13	Shortest_Distance_to_School	0.123984
14	Shortest_Distance_to_Subway	0.121152
8	LC_index	0.106573
18	YearMonth	0.102640
3	Crime_Rates	0.093738
2	CA_index	0.052059
5	HSP_index	0.051452
6	IR	0.041633
16	TC_index	0.039424
11	SDT_index	0.038430
4	Floor	0.022902
1	Building_Use	0.020333
17	UR	0.005931

# 부동산 전세가격 예측·전세가율 분석

## 로지스틱 회귀분석



## 로지스틱 회귀분석

가설 설정

- ⇒ 귀무가설 ( $H_0$ ) : 독립변수(매매가, 임대면적,... 금리)가 종속변수(전세가)에 **영향이 없다**.
- ⇒ 대립가설 ( $H_1$ ) : 독립변수가 종속변수(전세가)에 **영향이 있다**.
  - ▶ 독립변수 중 적어도 하나는 종속변수에 영향이 있다.



가설 검정을 위한 MNLogit 모델 생성

모형의 진단

- ⇒ 교차검증
- ⇒ 학습곡선

모형의 성능 - 부록 기재

- ⇒ Accuracy
- ⇒ F-1 score
- ⇒ McFadden's Pseudo  $R^2$

변수 선택

- ⇒ VIF
- ⇒ Wald 검정의 p-value

# 부동산 전세가격 예측·전세가율 분석

## 모형의 진단

⇒ 교차검증 - data leakage 여부 파악

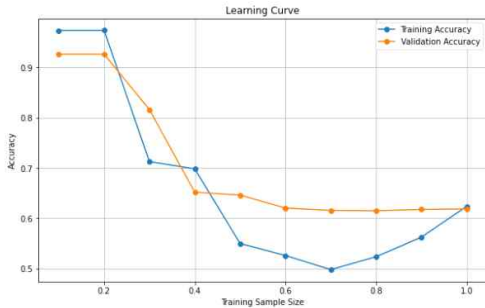
폴드	폴드별정확도	평균정확도
1	0.595264467	0.595206678
2	0.595778152	
3	0.595023678	
4	0.595216309	
5	0.594750783	



폴드 별 비슷한 정확도

Data Leakage 가능성 ↓

⇒ 학습곡선 - 모델의 과적합 감지



훈련 정확도가 과도하게  
증가하는 현상 없음

과적합 가능성 ↓

# 부동산 전세가격 예측·전세가율 분석

## 변수 선택

아파트

오피스텔

연립다세대

1차 제거 VIF>10

독립변수	VIF
TC_index	86.497
CA_index	79.168

독립변수	VIF
CA_index	48.400
TC_index	47.602

독립변수	VIF
CA_index	97.289
TC_index	89.541

2차 제거 p-value>0.01

전세가 Class	독립변수	p (유의확률)
1	HSP_index	0.247
	UR	0.506

전세가 Class	독립변수	p (유의확률)
1	IR	0.083

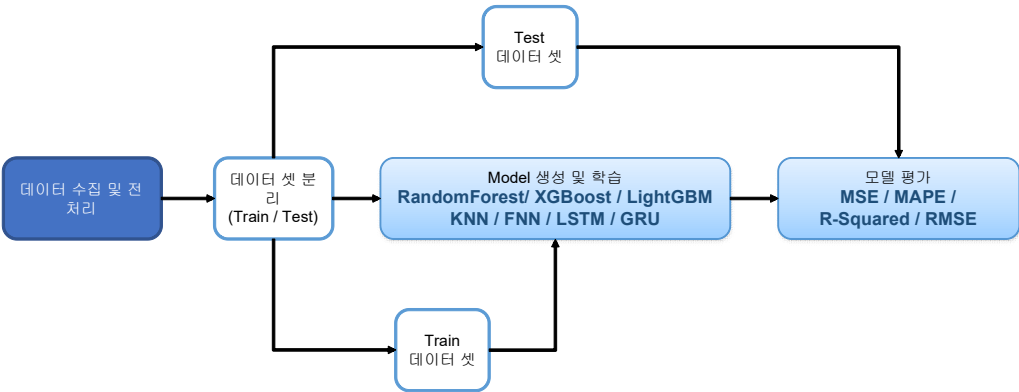
UR, HSP\_index, TC\_index, CA\_index 제거

IR, TC\_index, CA\_index 제거

TC\_index, CA\_index 제거

# 부동산 전세가격 예측·전세가율 분석

## 모델 훈련 및 평가 순서도





# 부동산 전세가격 예측·전세가율 분석

## 모델 생성 및 학습

### 머신 러닝

#### 지도 학습

KNN

#### 앙상블 모델

Random  
Forest

LightGBM

XGBoost

### 딥 러닝

FNN

LSTM

GRU

## 모델 생성 및 학습

### Random Forest

#### 배깅(Bagging)

여러 개의 의사 결정 트리를 구성하고 각각의 의사 결정 트리는 데이터를 기반으로 수행하고, 그 예측 결과를 모아서 최종 예측을 만듦

### XGBoost

그래디언트 부스팅(Gradient Boosting) 알고리즘을 기반으로 손실 함수의 그래디언트(기울기)를 활용하여 모델을 학습  
불필요한 트리 노드를 가지치기(Pruning)하고 트리의 깊이를 제한하여 모델을 단순화

### LightGBM

- Leaf-Wise 트리 분할 방식을 사용하여 최적 분할을 빠르게 찾고, 노은 예측 성능을 제공
- 카테고리형 데이터를 효율적으로 다룰 수 있어, 원-핫 인코딩(One-hot encoding) 없이도 카테고리 특성을 활용 가능

### 앙상블 모델 (Ensemble Model)

여러 개의 기본 모델을 조합하여 더 강력하고 정확한 예측 모델을 만드는 머신 러닝 기법

다수결 원칙에 기반으로 여러 개의 서로 다른 모델을 합쳐서 예측을 수행하고, 각 모델의 예측 결과를 조합하여 최종 예측을 만듦

보팅(Voting), 배깅(Bagging), 부스팅(Boosting), 스택킹(Stacking) 등의 예측 결합 방법을 사용

## 모델 생성 및 학습

### KNN

근접 이웃 기반으로 주어진 데이터 포인트 주변의 가장 가까운 이웃 데이터 포인트들을 사용하여 예측을 수행

데이터 포인트 간의 거리를 계산하여 이웃을 찾음

모델을 특정 함수 또는 파라미터로 피팅하지 않는 비모수적(non-parametric) 알고리즘

### LSTM

### FNN

# 부동산 전세가격 예측·전세가율 분석

모델 평가지표

## MAE

실제 값과 예측 값 간의 절대 오차를 나타내는 지표

각 데이터 포인트에서 실제 값과 예측 값의 차이를 모두 절대값으로 측정하고, 이 절대값 오차를 모든 데이터 포인트에 대해 평균을 낸 것

MAE가 작을수록 모델의 예측이 더 정확

$$MAE = \frac{\sum |y - \hat{y}|}{n}$$

## RMSE

오차 제곱값의 평균을 측정하고 루트를 씌운 지표

예측 오차를 평균화하면서 큰 오차에 패널티를 부여하는 특성을 가짐

RMSE가 작을수록 모델의 예측이 더 정확

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y - \hat{y})^2}{n}}$$

## MAPE

예측 오차를 백분율로 표현한 지표

각 데이터 포인트에서 실제 값과 예측 값의 백분율 오차를 계산하고, 이를 모든 데이터 포인트에 대해 평균을 낸 것

MAPE는 예측 성능을 백분율로 이해하기 용이

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y - \hat{y}}{y} \right|$$

## R-Squared

회귀 모델이 종속 변수의 분산을 얼마나 잘 설명하는지 측정하는 지표

0에서 1사이의 범위를 가지며, 1에 가까울수록 모델이 종속 변수의 분산을 잘 설명한다는 의미

R-Squared가 1에 가까울수록 모델이 더 좋은 성능을 가짐

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

## 성능 평가

	KNN	Random Forest	LGBR	XGBR	FNN	LSTM	GRU
MAE	6279.989	4006.445	3654.68	3760.784	4232.778	3886.538	4018.997
RMSE	9522.973	6645.358	6053.86	7135.28	6623.921	6223.073	6502.211
MAPE	0.218	0.139	0.128	0.166	0.147	0.132	0.139
R-Squared	0.752	0.879	0.899	0.87	0.879	0.893	0.887
Rank	7	5	1	2	6	3	4

## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)

주제 선정 배경

프로젝트 수행 방향

프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집

데이터 전처리

EDA

통계 분석

변수 설정

분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가

서울 자치구별 전세가격 예측

Web 구축 & 배포

### 04. 최종 결론

기대 효과

한계점

### 05. Document

데이터 정의서

참고문헌

출처









## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)

주제 선정 배경

프로젝트 수행 방향

프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집

데이터 전처리

EDA

통계 분석

변수 설정

분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가

서울 자치구별 전세가격 예측

Web 구축 & 배포

### 04. 최종 결론

기대 효과

한계점

### 05. Document

데이터 정의서

참고문헌

출처





## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)

주제 선정 배경

프로젝트 수행 방향

프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집

데이터 전처리

EDA

통계 분석

변수 설정

분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가

서울 자치구별 전세가격 예측

Web 구축 & 배포

### 04. 최종 결론

기대 효과

한계점

### 05. Document

데이터 정의서

참고문헌

출처

## [데이터 정의서]

Num	Column Name	Column Name(KOR)	Dtype	etc
1	Year	년		
2	Month	월		
3	Region_Code			
4	Region_Name			
5	KA_Code			
6	KA_Name			
7	LNA_Code			
8	LNA_Name			

## [데이터 정의서]

Num	Column Name	Column Name(KOR)	Dtype	etc
9	Main_num			
10	Sub_num			
11	Floor			
12	Building_Age			
13	Building_Use			
14	JS_Price			
15	JS_BA			
16	JS_PPA		float64	

## [데이터 정의서]

Num	Column Name	Column Name(KOR)	Dtype	etc
17	JS_PPP		float64	
18	address			
19	Num_hb		float64	
20	Population			
21	IR		float64	
22	UR		float64	
23	LC_index		float64	
24	CA_index		float64	



## [데이터 정의서]

Num	Column Name	Column Name(KOR)	Dtype	etc
25	TC_index		float64	
26	SDT_index		float64	
27	HSP_index		float64	
28	Crime_Rates		float64	
29	latitude		float64	
30	longitude		float64	
31	Sell_Price_mean	평균 매매 가격	float64	년도별, 월별, 구별, 유형별, 주소별 평균 매매 가격

## 모형의 성능

⇒ 아파트 데이터에 대한 정확도, F1-score,  $R^2$

평가지표	정확도	F1-score	McFadden's Pseudo $R^2$
지수	0.63075	0.56866	0.37286

⇒ 오피스텔 데이터에 대한 정확도, F1-score,  $R^2$

평가지표	정확도	F1-score	McFadden's Pseudo $R^2$
지수	0.39132	0.38637	0.14111

⇒ 엽립다세대 데이터에 대한 정확도, F1-score,  $R^2$

평가지표	정확도	F1-score	McFadden's Pseudo $R^2$
지수	0.68601	0.58466	NaN



Features\_Ranking의 상위 10가지 변수를 제외하고 만든 모델이기 때문에 낮은 성능을 보임

# 부동산 전세가격 예측·전세가율 분석

## 오즈비

$$\begin{aligned}\checkmark \text{ Odds} &= \frac{\text{전세가 클래스에 해당}}{\text{전세가 클래스에 해당하지 않음}} \\ &= \frac{p}{1-p} = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_i X_i}\end{aligned}$$

$$\text{오즈비 (OR)} = \frac{\frac{A}{B}}{\frac{C}{D}} = \frac{AD}{BC}$$

구분		전 세가클래스(예측)		
		YES	NO	합계
전 세가 클래스 (실제)	YES	A	B	A+B
	NO	C	D	C+D
합계		A+C	B+D	A+B C+D

- Odds = 해당 사건이 발생할 확률 / 발생하지 않을 확률
- OR>1 : 사건 발생 확률이 오즈비만큼 더 크게 나타남
- OR<1 : 사건 발생 확률이 오즈비만큼 더 작게 나타남
- OR=1 : 1에 가까울수록 사건이 발생할 확률에 유의미한 영향을 준다고 볼 수 없다고 판단

# 부동산 전세가격 예측·전세가율 분석

## 통계 분석

APT 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
1	Population	-0.0418	0.007	-6.393	0	0.959	-0.059	-0.025
	Crime_Rates	-0.1136	0.007	-17.465	0	0.893	-0.13	-0.097
	YearMonth_encoded	0.942	0.014	68.842	0	2.565	0.907	0.977
	HSP_index	-0.016	0.014	-1.157	0.247	0.984	-0.052	0.02
	SDT_index	-0.1018	0.011	-9.413	0	0.903	-0.13	-0.074
	IR	-0.1471	0.013	-11.724	0	0.863	-0.179	-0.115
	UR	-0.0051	0.008	-0.666	0.506	0.995	-0.025	0.015
	Floor	0.0202	0.007	3.02	0.003	1.02	0.003	0.037
2	Population	-0.1233	0.007	-17.351	0	0.884	-0.142	-0.105
	Crime_Rates	-0.2613	0.007	-35.716	0	0.77	-0.28	-0.242
	YearMonth_encoded	1.6035	0.016	102.374	0	4.97	1.563	1.644

## 통계 분석

APT 전세가별 독립변수들의 통계량

전세가 Class	독립 변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
2	HSP_index	-0.2258	0.015	-15.209	0	0.798	-0.264	-0.188
	SDT_index	-0.388	0.012	-32.641	0	0.678	-0.419	-0.357
	IR	-0.7147	0.014	-49.307	0	0.489	-0.752	-0.677
	UR	-0.4233	0.008	-50.491	0	0.655	-0.445	-0.402
	Floor	0.089	0.007	12.042	0	1.093	0.07	0.108
3	Population	0.1161	0.008	14.81	0	1.123	0.096	0.136
	Crime_Rates	-1.079	0.011	-97.512	0	0.34	-1.108	-1.051
	YearMonth_encoded	2.9254	0.018	161.788	0	18.641	2.879	2.972
	HSP_index	-0.5844	0.016	-35.723	0	0.557	-0.627	-0.542
	SDT_index	0.1711	0.015	11.508	0	1.187	0.133	0.209
	IR	0.4112	0.016	26.41	0	1.509	0.371	0.451

## 통계 분석

APT 전세가별 독립변수들의 통계량

전세가 Class	독립 변수	회귀 계수	S.E.	Wald	p (유의 확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
3	UR	-0.2894	0.009	-31.192	0	0.749	-0.313	-0.266
	Floor	-0.727	0.009	-77.312	0	0.483	-0.751	-0.703
4	Population	-2.9085	0.02	-143.339	0	0.003	-2.961	-2.856
	Crime_Rates	-3.1043	0.023	-136.49	0	0.055	-3.163	-3.046
	YearMonth_encoded	4.6526	0.04	115.655	0	0.045	4.549	4.756
	HSP_index	-0.8101	0.039	-20.993	0	104.855	-0.909	-0.711
	SDT_index	-5.0977	0.038	-132.662	0	0.445	-5.197	-4.999
	IR	-6.3749	0.051	-125.013	0	0.006	-6.506	-6.244
	UR	-3.6309	0.026	-137.941	0	0.002	-3.699	-3.563
	Floor	1.4547	0.016	89.271	0	0.026	1.413	1.497

# 부동산 전세가격 예측·전세가율 분석

## 통계 분석

Officetel 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
1	Population	-0.1366	0.007	-18.611	0	0.872	-0.156	-0.118
	Crime_Rates	0.0291	0.007	4.11	0	1.03	0.011	0.047
	YearMonth_encoded	0.4018	0.009	44.858	0	1.495	0.379	0.425
	HSP_index	-0.4685	0.008	-55.121	0	0.626	-0.49	-0.447
	SDT_index	0.088	0.009	9.475	0	1.092	0.064	0.112
	IR	-0.0199	0.011	-1.735	0.083	0.98	-0.05	0.01
	UR	-0.0391	0.008	-5.167	0	0.962	-0.059	-0.02
	Floor	0.0414	0.006	6.397	0	1.042	0.025	0.058
2	Population	-0.252	0.008	-33.144	0	0.777	-0.272	-0.232
	Crime_Rates	-0.2329	0.008	-28.603	0	0.792	-0.254	-0.212
	YearMonth_encoded	0.6472	0.01	61.978	0	1.91	0.62	0.674

# 부동산 전세가격 예측·전세가율 분석

## 통계 분석

Officetel 전세가별 독립변수들의 통계량

전세가 Class	독립 변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
2	HSP_index	-1.0736	0.01	-107.96	0	0.342	-1.099	-1.048
	SDT_index	-0.2494	0.01	-25.804	0	0.779	-0.274	-0.224
	IR	-0.7263	0.014	-51.548	0	0.484	-0.763	-0.69
	UR	-0.0367	0.008	-4.599	0	0.964	-0.057	-0.016
	Floor	-0.0858	0.007	-12.129	0	0.918	-0.104	-0.068
3	Population	0.2088	0.008	27.218	0	1.232	0.189	0.229
	Crime_Rates	-0.0909	0.009	-9.943	0	0.913	-0.114	-0.067
	YearMonth_encoded	1.0304	0.011	92.489	0	2.802	1.002	1.059
	HSP_index	-0.3933	0.009	-43.591	0	0.675	-0.416	-0.37
	SDT_index	0.0395	0.01	3.903	0	1.04	0.013	0.066
	IR	-0.2354	0.013	-17.718	0	0.79	-0.27	-0.201



# 부동산 전세가격 예측·전세가율 분석

## 통계 분석

Officetel 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
3	UR	0.1105	0.008	14.271	0	1.117	0.091	0.13
	Floor	0.0197	0.007	2.892	0.004	1.02	0.002	0.037
4	Population	-0.202	0.009	-22.203	0	0.817	-0.225	-0.179
	Crime_Rates	-1.3403	0.016	-84.006	0	0.262	-1.381	-1.299
	YearMonth_encoded	1.7522	0.015	120.505	0	5.767	1.715	1.79
	HSP_index	-0.7354	0.013	-58.641	0	0.479	-0.768	-0.703
	SDT_index	0.377	0.013	28.665	0	1.458	0.343	0.411
	IR	-0.2796	0.02	-13.64	0	0.756	-0.332	-0.227
	UR	0.4174	0.009	45.522	0	1.518	0.394	0.441
	Floor	-0.5651	0.008	-67.73	0	0.568	-0.587	-0.544

# 부동산 전세가격 예측·전세가율 분석

## 통계 분석

Townhouse 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
1	Population	0.6509	0.01	65.185	0	1.917	0.625	0.677
	Crime_Rates	0.4193	0.01	44.092	0	1.521	0.395	0.444
	YearMonth_encoded	1.7948	0.023	76.829	0	6.018	1.735	1.855
	HSP_index	-1.0262	0.023	-45.488	0	0.358	-1.084	-0.968
	SDT_index	0.2602	0.013	19.897	0	1.297	0.227	0.294
	IR	0.5075	0.014	35.172	0	1.661	0.47	0.545
	UR	0.3526	0.009	37.37	0	1.423	0.328	0.377
	Floor	0.0725	0.008	9.024	0	1.075	0.052	0.093
2	Population	-7.1568	0.408	-17.539	0	0.001	-8.208	-6.106
	Crime_Rates	6.3095	0.161	39.301	0	549.753	5.896	6.723
	YearMonth_encoded	-20.6099	3.231	-6.379	0	0	-28.933	-12.287

## 통계 분석

Townhouse 전세가별 독립변수들의 통계량

전세가 Class	독립 변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
2	HSP_index	362.7142	17.218	21.065	0	3.35E+157	318.362	407.066
	SDT_index	31.5697	1.11	28.447	0	5.13531E+13	28.711	34.428
	IR	87.375	2.667	32.766	0	8.84E+37	80.506	94.244
	UR	12.3919	0.864	14.342	0	240841.239	10.166	14.617
	Floor	2.6607	0.135	19.763	0	14.307	2.314	3.008
4	Population	-0.6846	0.1	-6.851	0	0.504	-0.942	-0.427
	Crime_Rates	4.4764	0.121	37.011	0	87.917	4.165	4.788
	YearMonth_encoded	117.8526	3.575	32.968	0	1.52E+51	108.645	127.061
	HSP_index	-38.8477	1.708	-22.742	0	0	-43.248	-34.448
	SDT_index	23.8996	0.884	27.044	0	2.39587E+10	21.623	26.176
	IR	10.8796	1.064	10.222	0	53081.064	8.138	13.621

## 통계 분석

Townhouse 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
4	UR	10.8816	0.681	15.989	0	53190.779	9.129	12.635
	Floor	-12.2244	0.23	-53.041	0	0	-12.818	-11.631

## [참고문헌]

### 뉴스기사

- 윤종은(2023.07.18) “전세사기 막을 보완 대책 시급”(오마이뉴스) <https://omn.kr/24uhx>
- 심민규(2023.08.02) 무자본 갭투자로 126채 산 뒤 전세보증금 250억 편취한 일당(연합뉴스) <https://www.yna.co.kr/view/AKR20230802045200060>
- 박형기(2023.08.28) 파월은 물론 연준 간부들 잇달아 추가 금리인상 시사(뉴스1) <https://www.news1.kr/articles/5152515>
- 우경희(2023.08.29) 中 부동산위기 전망지 높다, 상반기만 6조 손실...“주식 거래재개”(머니투데이) <https://news.mt.co.kr/mtview.php?no=2023082809170892576>

### 논문

- 김학현, 유한규, and 오하영. (2023). 딥러닝과 머신러닝을 이용한 아파트 실거래가 예측. 정보처리학회논문지. 소프트웨어 및 데이터 공학, 12(2), 59-76.
- Yadav, S., Dhanda, N., Sahai, A., Verma, R., Pandey, S. (2023). Real Estate Price Prediction Using Machine Learning. In: Sarkar, D.K., Sadhu, P.K., Bhunia, S., Samanta, J., Paul, S. (eds) Proceedings of the 4th International Conference on Communication, Devices and Computing. ICCDC 2023. Lecture Notes in Electrical Engineering, vol 1046. Springer, Singapore. [https://doi.org/10.1007/978-981-99-2710-4\\_9](https://doi.org/10.1007/978-981-99-2710-4_9)
- 고주형 and 강명구. (2019). 부동산 가격 요인과 가격상승률 요인 비교 연구: 서울시 재건축 아파트를 중심으로. 부동산학연구, 25(2), 7-22.
- 최남진. (2023). 초과 유동성과 전국 아파트 가격 간의 관계에 대한 연구. 부동산분석, 9(1), 195-210.
- 김은미. (2022). 거시경제변수가 지역 별 아파트 전세가격에 미치는 영향 및 예측모델 구축에 관한 연구. 지적과 국토정보, 52(2), 211-231.
- M Sri Rangan, W. (2023) *Creating accurate valuation models for real estate properties.*
- Prof. Dr. Sven Bienrt 등 16명 (2023) *Real Estate Valuation in the Age of Artificial Intelligence - Modern Machine Learning Algorithms and their Application in Property Appraisal*

## [출처]

- 남녀 픽토그램 [https://kr.freepik.com/free-vector/flat-design-male-female-symbols\\_22340731.htm#query=%EB%82%A8%EB%85%80%20%ED%94%BD%ED%86%A0%EA%B7%B8%EB%9E%A8&position=15&from\\_view=keyword&track=ais](https://kr.freepik.com/free-vector/flat-design-male-female-symbols_22340731.htm#query=%EB%82%A8%EB%85%80%20%ED%94%BD%ED%86%A0%EA%B7%B8%EB%9E%A8&position=15&from_view=keyword&track=ais)
- ppt 템플릿 <http://pptbizcam.co.kr/?p=6897>
- 서울 전세 실거래: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/OA-21276/S/1/datasetView.do>)
- 서울 매매 실거래: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/OA-21275/S/1/datasetView.do>)
- 기준 금리: 한국은행 경제통계시스템 (<https://ecos.bok.or.kr/#/SearchStat>)
- 실업률: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/59/S/2/datasetView.do>)
- 전세수급동향:
- 지하철역 위치: 레일포털 (<https://data.kric.go.kr/rips/>)
- 학교(초, 중, 고등학교) 위치: 학구도안내서비스 (<https://schoolzone.emac.kr/publicData/publicDataView.do>)
- 대학 위치: 공공데이터포털 (<https://www.data.go.kr/data/15052369/fileData.do>)
- 서울 공원 위치: 서울 열린 데이터 광장 (<https://www.data.go.kr/data/15051894/fileData.do>)
- 인구수: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/10043/S/2/datasetView.do?stcSrl=10043>)
- 서울 5대 범죄: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/316/S/2/datasetView.do>)
- 경기종합지수: 국가통계포털  
([https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT\\_1C80158vw\\_cd=MT\\_ZTITLE&list\\_id=J1\\_1&seqNo=&lang\\_mode=ko&language=kor&obj\\_var\\_id=&itm\\_id=&conn\\_path=MT\\_ZTITLE](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1C80158vw_cd=MT_ZTITLE&list_id=J1_1&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE))
- 서울 주택 재개발: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/241/S/2/datasetView.do>)