



# 전세영향변수를 통한 역전세난 생존분석

시계열자료분석팀

김민우 김영호 정승연 조건우 조웅빈





# INDEX



분석 배경



시계열 클러스터링



변수선택을 위한 모델링



생존분석



정책 제언 및 한계와 의의

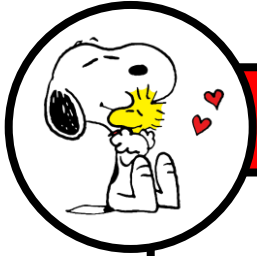




분석배경



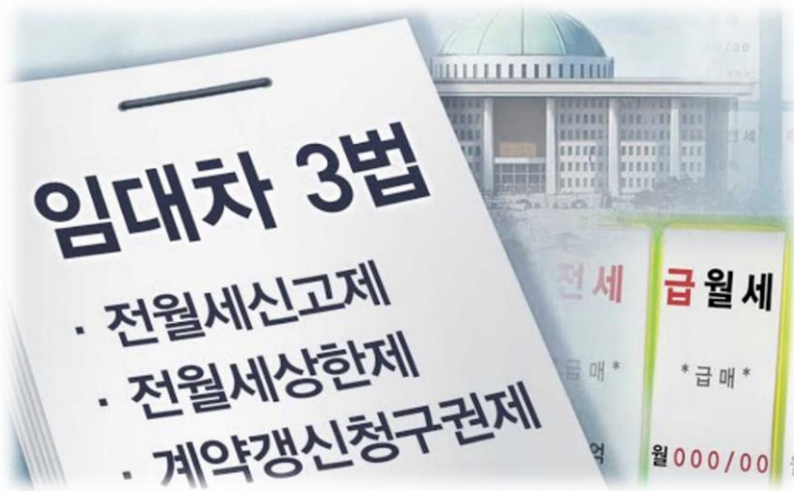
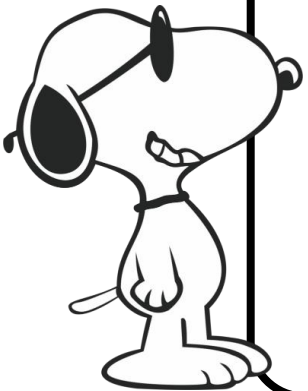
## 분석배경



### 임대차3법

: 2020년 7월 31일 시행된 주택임대차보호법 개정안

구체적으로  
살펴보아야계군!



계약 갱신 청구권

전월세 상한제

전월세 신고제



## 분석배경

### 계약 갱신 청구권

임차인이 희망하는 경우 1회 계약 갱신을 청구할 수 있는 권리로,  
임차인의 안심 거주기간이 2년 더 늘어나게 됨

### 전월세 상한제

계약 갱신 시 임대료 증액 상한을 5% 범위내로 제한하여  
임차인들의 임대료 급등으로 인한 부담을 줄이기 위한 목적으로 시행

### 전월세 신고제

세입자의 거주사실을 증명해주고 보호해주기 위해 도입된 정책으로,  
신고 시 자동으로 확정일자를 부여 받음



## 분석배경



현 상황에서의 문제점은 임대차법 시행 2년을 맞아  
계약 갱신 청구권 만료로 부동산 시장에 물량이  
쏟아지면서 혼란이 야기되고 있다는 점이야!



## 분석배경

### 서울 아파트값 하락



뉴스1 PICK 2시간 전 | 네이버뉴스

"신고가 행진→하락가 행진"...서울 아파트 10건 중 7건 '하락거래'

지난해까지만 해도 10건 중 7건이 최고가에 거래됐던 서울 아파트 시장 상황도 반전됐다. 지난달에는... 신고가보다 낮은 가격에 팔린 하락 거래 매물은 제외한다. ...



한겨레 PICK 1일 전 | 네이버뉴스

2~3억씩 뚝 뚝...'노도강' 아파트값 가파른 추락, 왜

그러나 노원구는 -4.38%, 도봉구는 -4.28%, 강북구는 -3.15% 등 '노도강' 아파트값 하락 폭은 서울 평균치의 갑절 수준에 이른다. 이에 반해 고가주택이 밀집해 있는...



특히 서울 아파트값을 살펴보면, 9년여 만에 가장 큰 폭으로 떨어졌고,  
전세금도 13주 연속 하락세를 보임



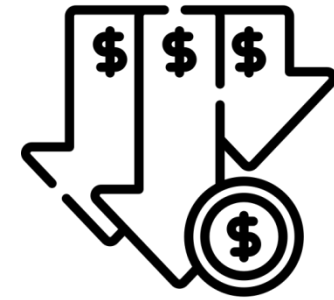
## 분석배경



대출금리 상승으로  
반전세·월세 선호가  
지속



신규 전세  
수요 감소



매매가격 하락 +  
전세매물 가격  
하향 조정





## 분석배경

이 때문에 비강남권은 물론 고액 전세가 많은 강남권도 전체 물건이 적체되면서 집주인이 세입자에게 보증금을 돌려줘야 하는 역전세난이 본격화하고 있다.

서울 송파구 잠실동 리센츠 아파트 전용면적 84m<sup>2</sup>의 경우 현재 전세 물건이 12억원 선에 나온다는 게 현지 중개업소의 설명이다.

2020년 9~10월 이 아파트의 전세 계약 금액이 최고 13억~14억원이었는데 이보다 1억~2억원가량 낮은 것이다.

잠실 엘스 전용 84m<sup>2</sup>도 현재 시중에 나와 있는 전세물건의 시세가 11억~12억원 수준으로, 2년 전에 최고 12~14억원에 전세계약을 체결한 집주인은 세입자에게 1억원 이상의 보증금을 돌려줘야 한다.

**역전세난**이란 전세난에 반대되는 의미로 전세의 공급은 늘었으나 수요가 줄어들면서 전세계약이 이루어지지 않는 어려움을 의미





## 분석배경



서울 강남의 경우 시세보다 전셋값을 1억~2억 이상 낮춰야 계약이 가능하며, 집주인은 전세 만기가 임박해도 세입자를 구하지 못함



역전세난 현상으로 인해 만기 후 보증금을 돌려주지 못하기도 함





## 분석배경



임대차3법의 시행과 최근 나타난 역전세난 현상에

주목하여 향후 시장상황에 대한 관련 인사이트를 도출하고,

전세 제도에 대한 전반적인 흐름에 대해 살펴보기로 결정



서울 임대의 경우 전세보다 전셋값을 1억 ~ 2억 이상 낮춰야 계약이 가능하고, 집주인은 전세 만기▲가 임박해도 세입자를 구하지 못함



시계열, 머신러닝, 딥러닝 모델과 더불어 생존분석 모델을 활용하여



미래에 대한 인사이트 도출 및 정책 제언

역전세난 현상으로 인해 만기 후 보증금을 돌려주지 못하기도 함

토이 프로젝트를 통해 분석단위를 구에서 확장하여



구별 법정동에 따른 다양한 시계열 모델 적용



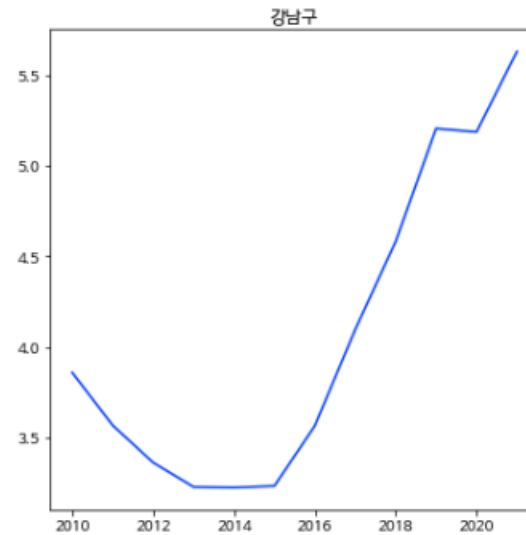
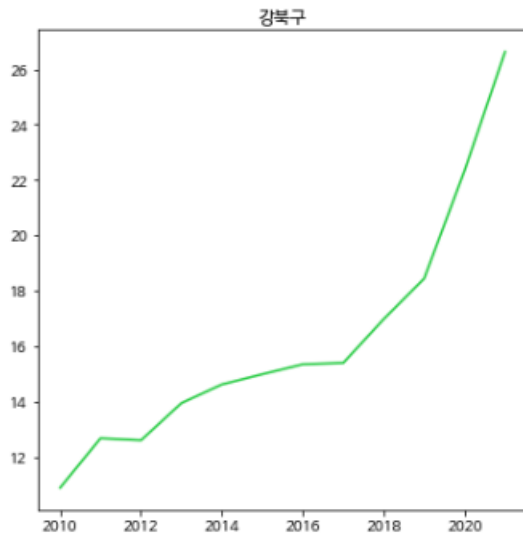


시계열 클러스터링



## 시계열 클러스터링

### 클러스터링의 필요성



구별로 시계열적 특성이 다양함

➡ 비슷한 구끼리 그룹으로 만들어 분석!



## 시계열 클러스터링

### 클러스터링 방법 결정

But 시계열의 자기상관성으로 인해 일반적인 K-means 클러스터링 사용 불가!



DTW를 활용한 클러스터링

시계열 클러스터링 방법으로 DTW를 기준으로 K-means 클러스터링



## 시계열 클러스터링

### 최적의 클러스터 개수 결정

#### Elbow method

클러스터 간의 거리의 합 계산

Inertia



그래프가 급격히 꺾이는 지점 k 찾음

*팔꿈치와 비슷한 형태를 보임*

#### 실루엣 계수

클러스터 개수별 실루엣 계수 계산



1에 가장 가까운 값을 이용하여

클러스터 개수 결정

*실루엣 계수는 (-1, 1)의 범위를 가짐*

#### DB Index

그룹 쌍 크기를 중심 간 거리로 나눔

*그룹 간 분포를 그룹 내 비율로 나눔*



가장 작은 값을 통해

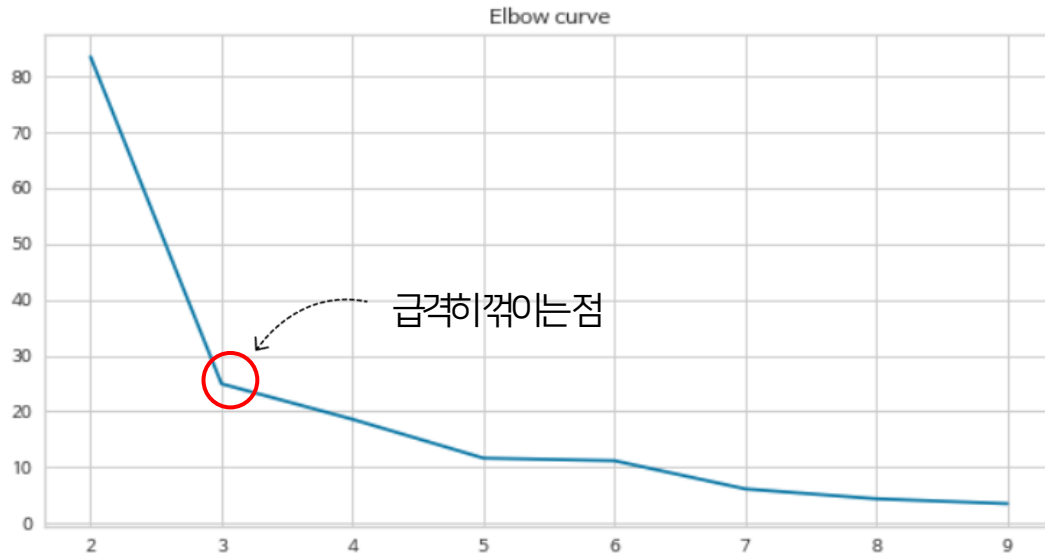
클러스터 개수 결정

*작을수록 그룹 간 분리가 잘 됨!*



## 시계열 클러스터링

### Elbow method



K=3 일때 Inertia 그래프가 급격히 꺾임



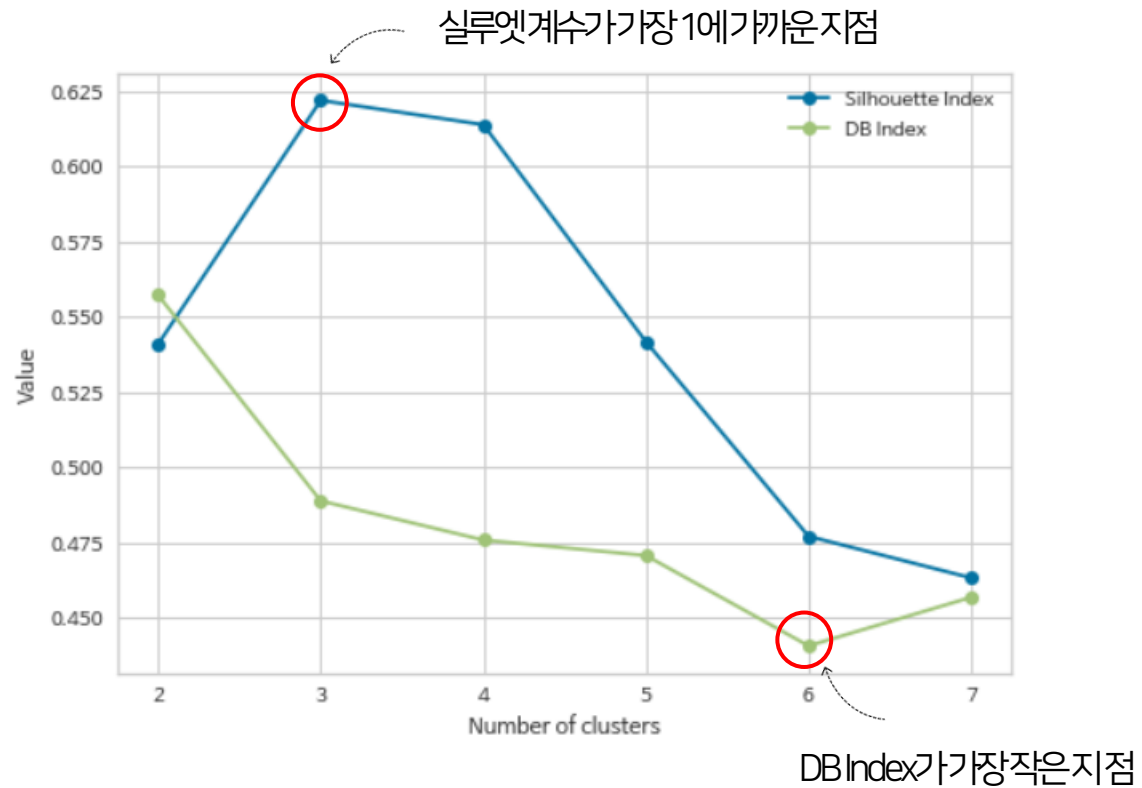
최적의 클러스터 개수 후보 **K=3**





## 시계열 클러스터링

### 실루엣계수 & DB index



K=3일 때 실루엣 계수가 가장 큼

&

K=6일 때 DB Index가 가장 작음



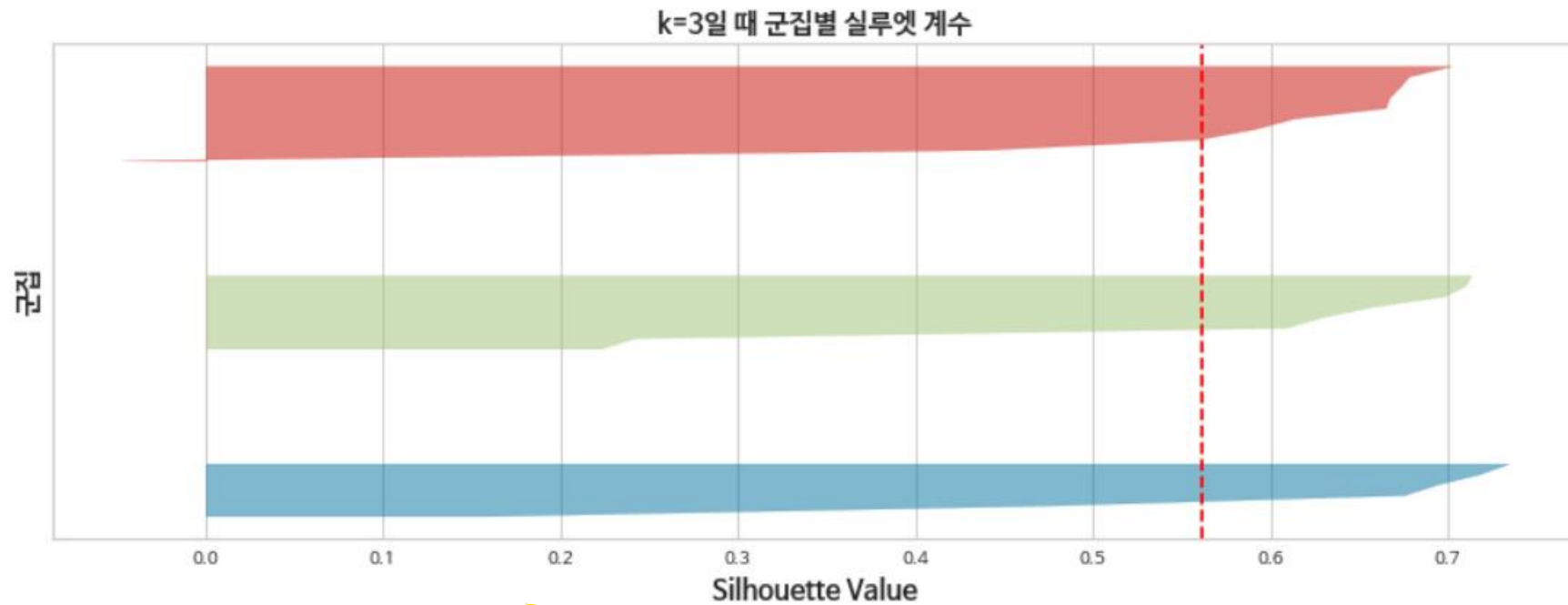
최종 클러스터 개수 후보

K=3 or K=6



## 시계열 클러스터링

### 최적의 클러스터 개수 결정



최적의 클러스터 개수 3개로 선정!

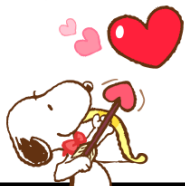
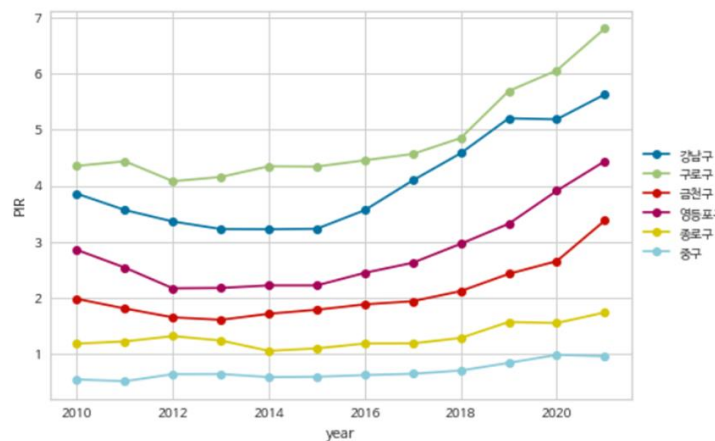
*elbow method, 실루엣 계수 & DB Index*  
모두 최적!

# 시계열 클러스터링

## 클러스터링 결과

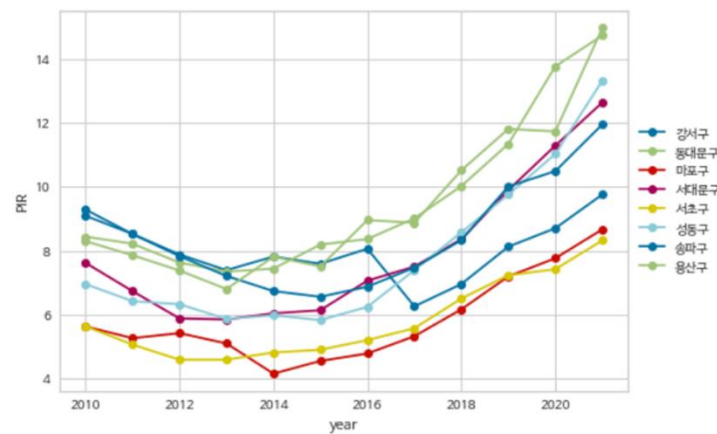
## 군집 1

완만한변화후상승



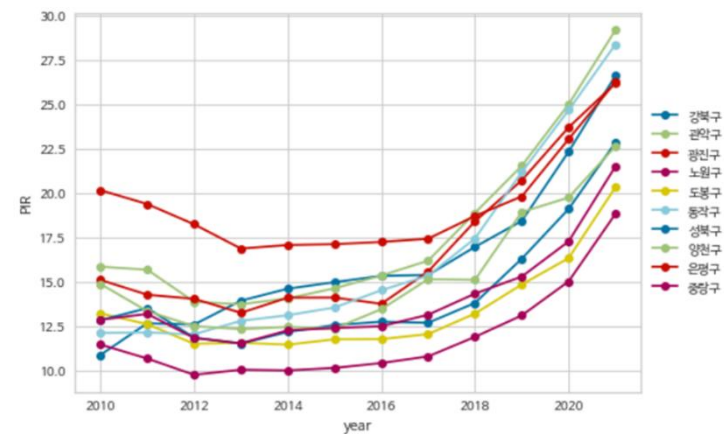
## 군집 2

하락후급상승



## 군집 3

완만한변화후급상승





**변수선택을 위한 모델링**



## 변수선택을 위한 모델링



TREE



XGBOOST

*Extreme Gradient  
Boosting Model*

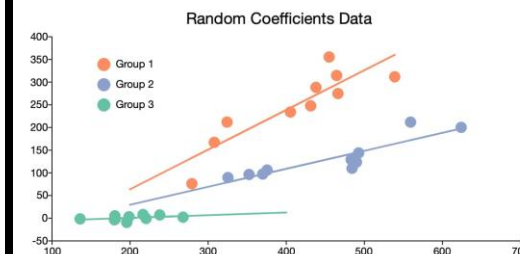
TREE



LGBM

*Light Gradient  
Boosting Model*

Panel



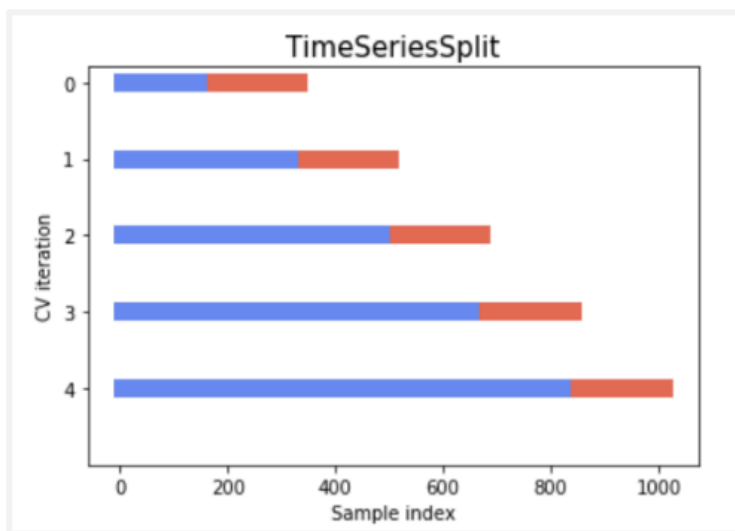
패널 모델

*Panel Model*



## 변수선택을 위한 모델링 : XGBOOST & LGBM

### Time-Series CV



시간 순서를 고려한 교차검증(Cross Validation)

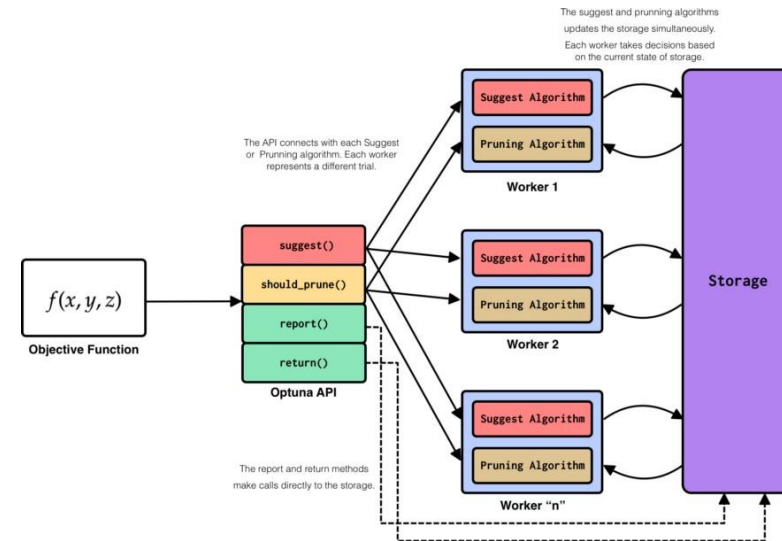
누적하며 이동하는 Window 내에서 train/validation set을 분리하는 기법

자세한 내용은 시계열팀 3주차 클린업 참고!



## 변수선택을 위한 모델링 : XGBOOST & LGBM

### OPTUNA



기계 학습(ML)을 위해 설계된 자동 하이퍼파라미터 최적화 soft framework

다른 튜닝 방법에 비해 처리 속도, 사용성, 문서화, 시각화 제공 등의 장점을 가짐



## 변수선택을 위한 모델링 : 패널모형

### 정규화 & 정상화

[min-max scaling]

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

[1차 차분(differencing)]

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t$$

변수 단위의 영향을 제거하기 위해 min-max scaling 진행

Linear한 추세를 제거하기 위해 1차 차분 진행





## 변수선택을 위한 모델링 : 패널모형

### 모형 검정

#### DW Test

자기상관 검정방법 중 하나

2 근처의 값이 나올 때

자기상관이 존재하지 않음!

0-2 : 양의 자기상관관계

2 : 자기상관관계가 없음

2-4 : 음의 자기상관관계

#### ADF

단위근 검정 방법 중 하나

$H_0$  : 비정상성을 만족한다.

$H_1$  : 정상성을 만족한다.

#### KPSS

단위근 검정 방법 중 하나

$H_0$  : 정상성을 만족한다.

$H_1$  : 비정상성을 만족한다.



## 변수선택을 위한 모델링 : 패널모형



### DW Test

자기상관 검정방법 중 하나

2 근처의 값이 나올 때

자기상관이 존재하지 않음!

0-2 : 양의 자기상관관계

2 : 자기상관관계가 없음

2-4 : 음의 자기상관관계

### ADF

#### 패널모형 사후 검정 결과

단위근 검정 방법 중 하나

DW Test와 ADF/KPSS 검정을 통해

최종적으로 자기상관 X, 정상성을 만족!

$H_0$  : 비정상성을 만족한다.

$H_1$  : 정상성을 만족한다.

### KPSS

단위근 검정 방법 중 하나

$H_0$  : 정상성을 만족한다.

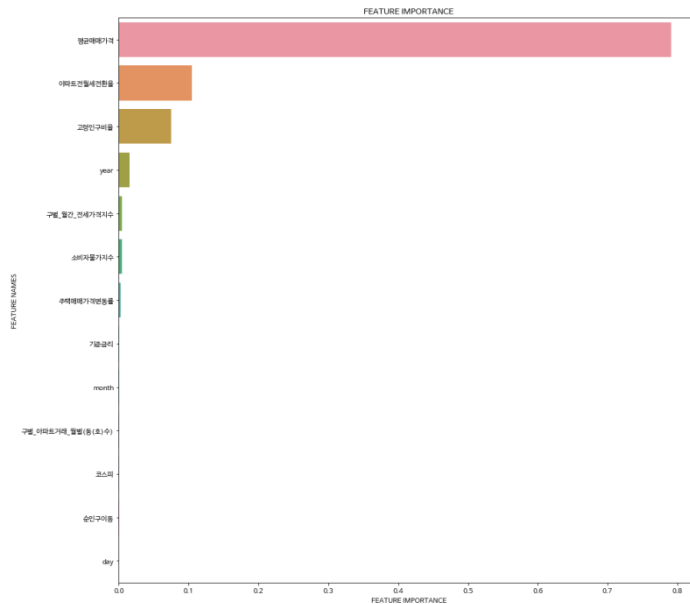
$H_1$  : 비정상성을 만족한다.



## 모델링 결과해석

### XGBOOST

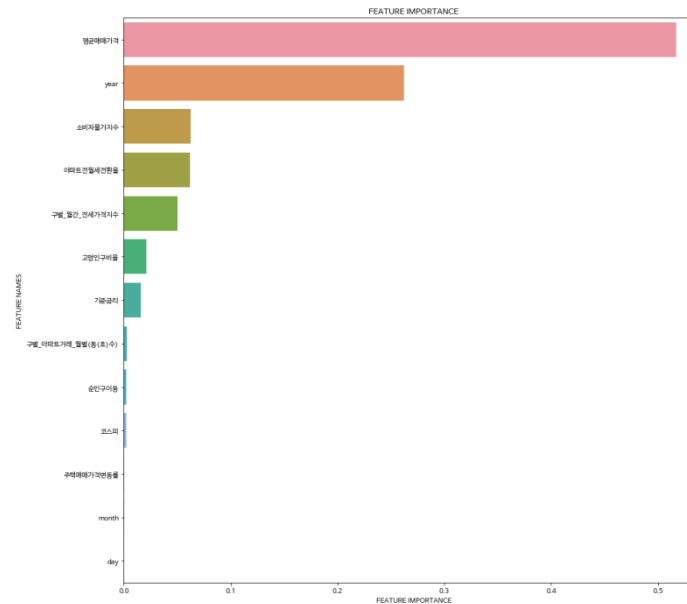
군집1



평균매매가격 > 아파트전월세전환율 >

고령인구비율 > ...

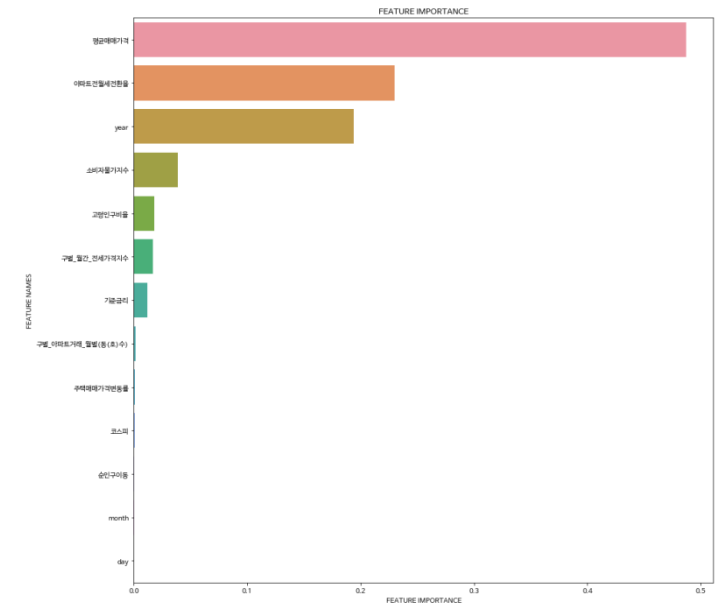
군집2



평균매매가격 > 소비자물가지수 >

아파트전월세전환율 > ...

군집3



평균매매가격 > 아파트전월세전환율 >

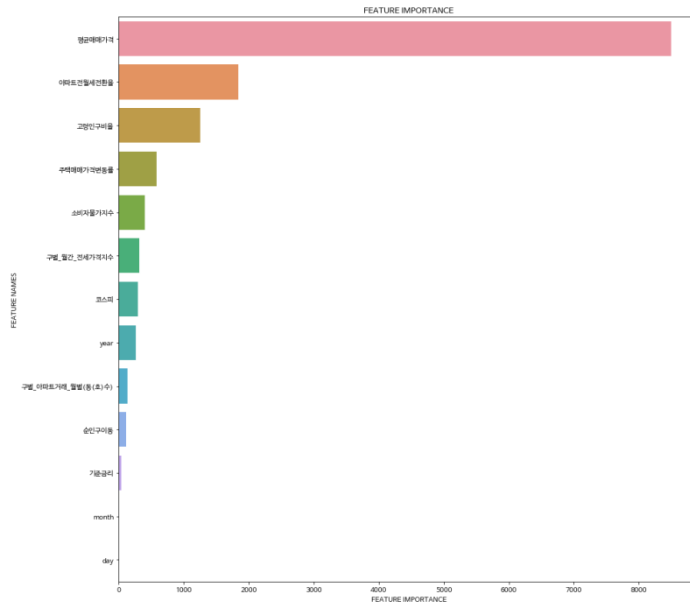
소비자물가지수 > ...



## 모델링 결과해석

### LGBM

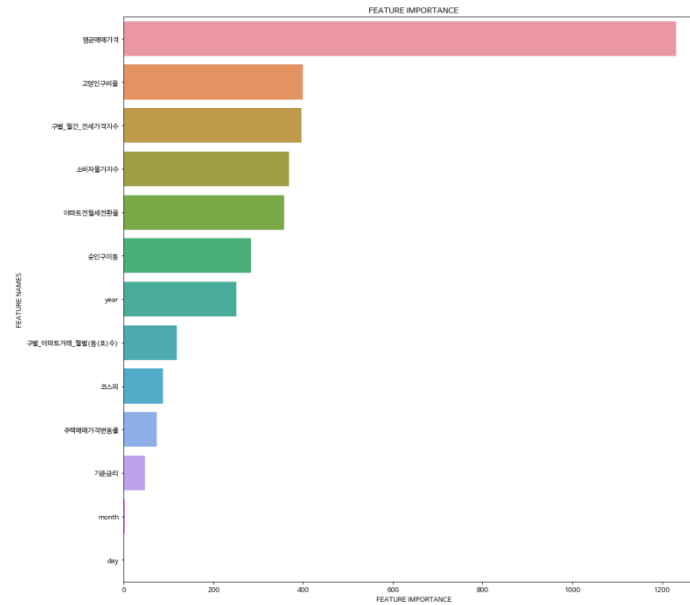
군집1



평균매매가격 > 아파트전월세전환율 >

고령인구비율 > ...

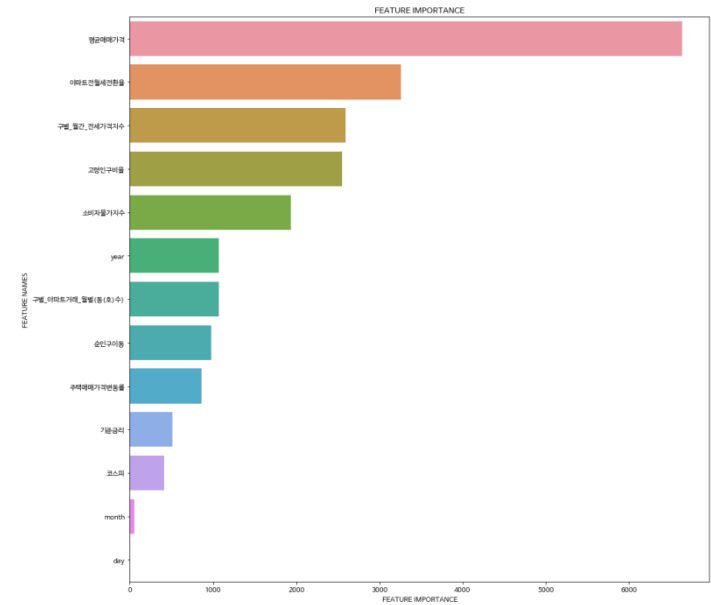
군집2



평균매매가격 > 고령인구비율 >

전세가격지수 > ...

군집3



평균매매가격 > 아파트전월세전환율 >

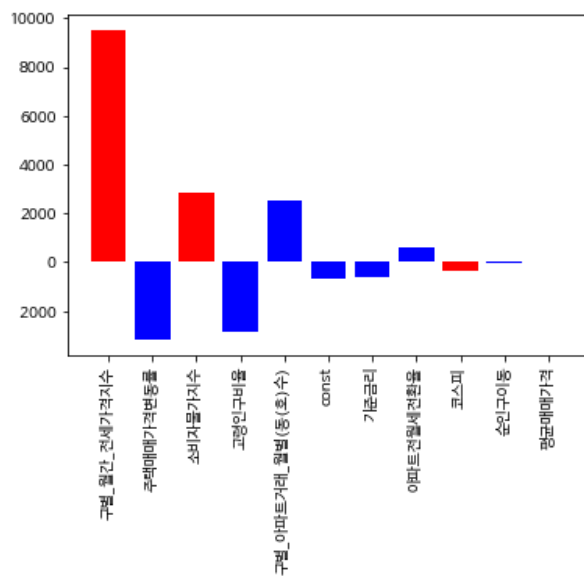
전세가격지수 > ...



## 모델링 결과해석

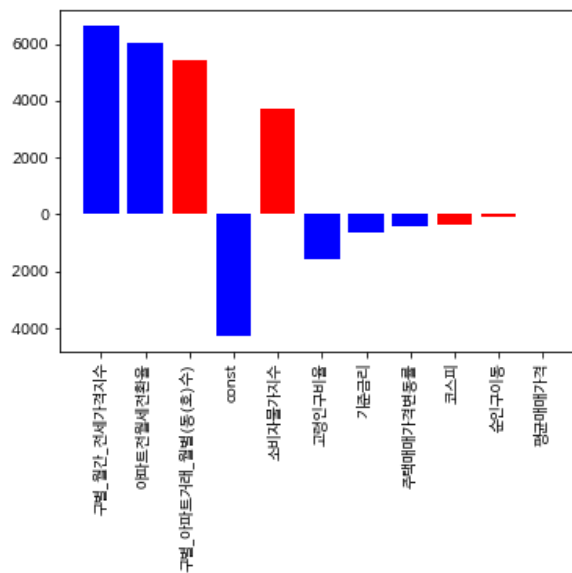
### 패널 모형

군집1



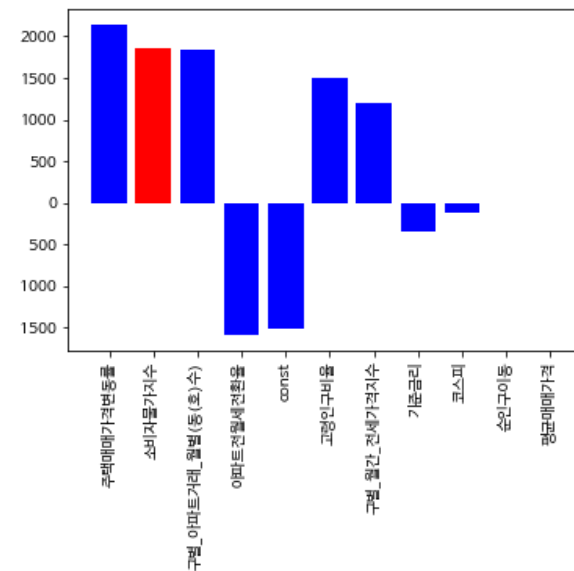
구별\_월간\_전세가격지수, 소비자물가  
지수, 코스피

군집2



구별\_아파트거래\_월별, 소비자물가지수,  
순인구이동, 코스피, 평균매매가격

군집3



소비자물가지수, 평균매매가격

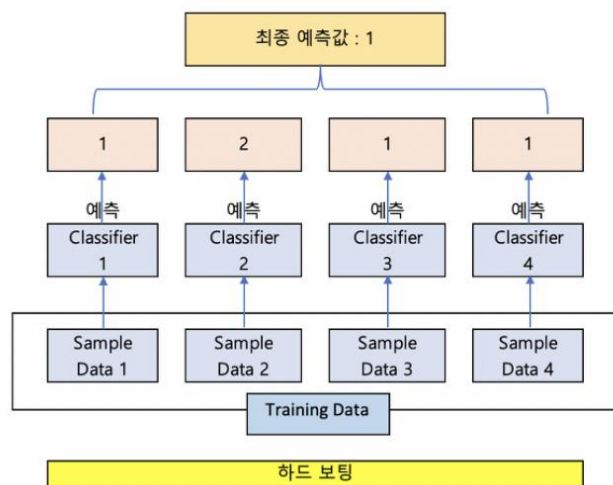


## 모델링 결과해석

### Voting

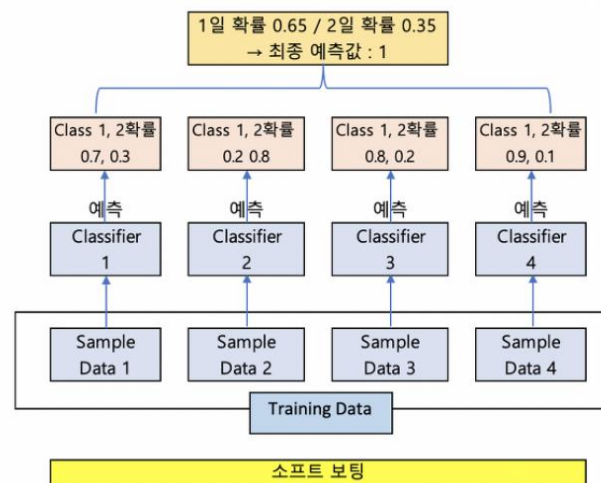


#### [Hard Voting]



다수의 추정기가 결정한 예측값들 중  
많은 Label을 선택

#### [soft Voting]



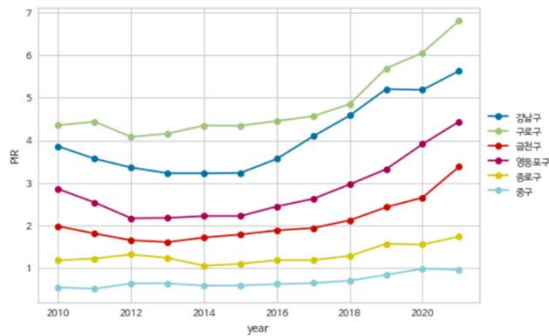
다수의 추정기가 예측한 확률들의 평균을 통해  
높은 확률의 Label을 선택



## 모델링 결과해석

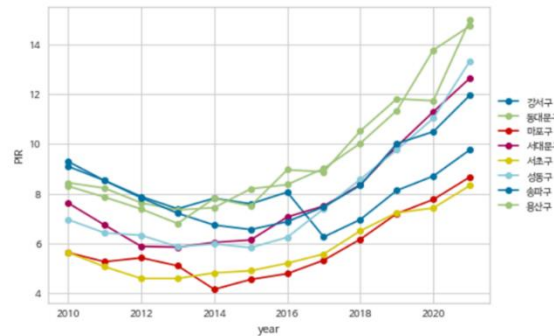
### 최종 중요 변수

군집1



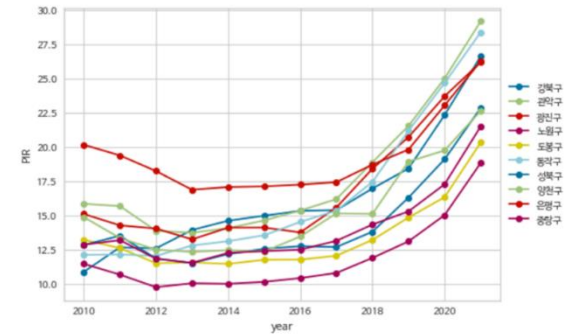
고령인구비율, 구별\_월간\_전세가격지수,  
소비자물가지수, 아파트전월세전환율,  
코스피, 평균매매가격

군집2



고령인구비율, 구별\_월간\_전세가격지수,  
구별\_아파트거래\_월별, 소비자물가지수,  
순인구이동, 아파트전월세전환율, 평균  
매매가격

군집3



고령인구비율, 구별\_월간\_전세가격지수,  
기준금리, 소비자물가지수, 아파트전월  
세전환율, 평균매매가격



생존분석

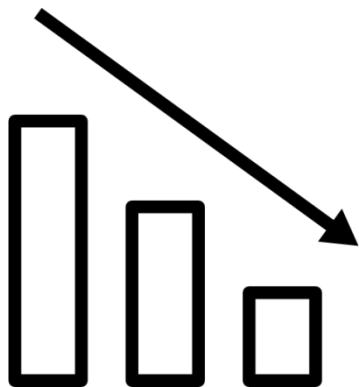




## 생존분석 : 모델링

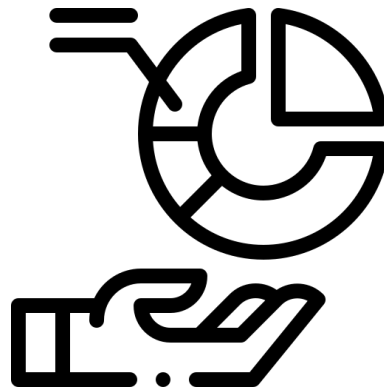
### 생존분석 모델

#### Non-Parametric(비모수적)



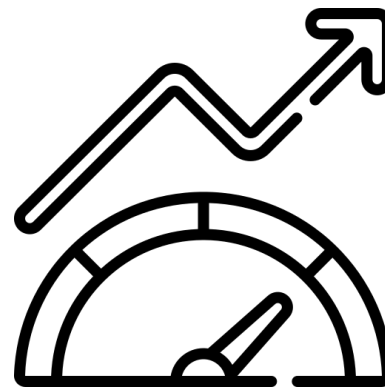
Kaplan-Meier  
Estimation(Curve)

#### Semi-Parametric(준모수적)



Cox Proportional  
Hazard model

#### Parametric(모수적)

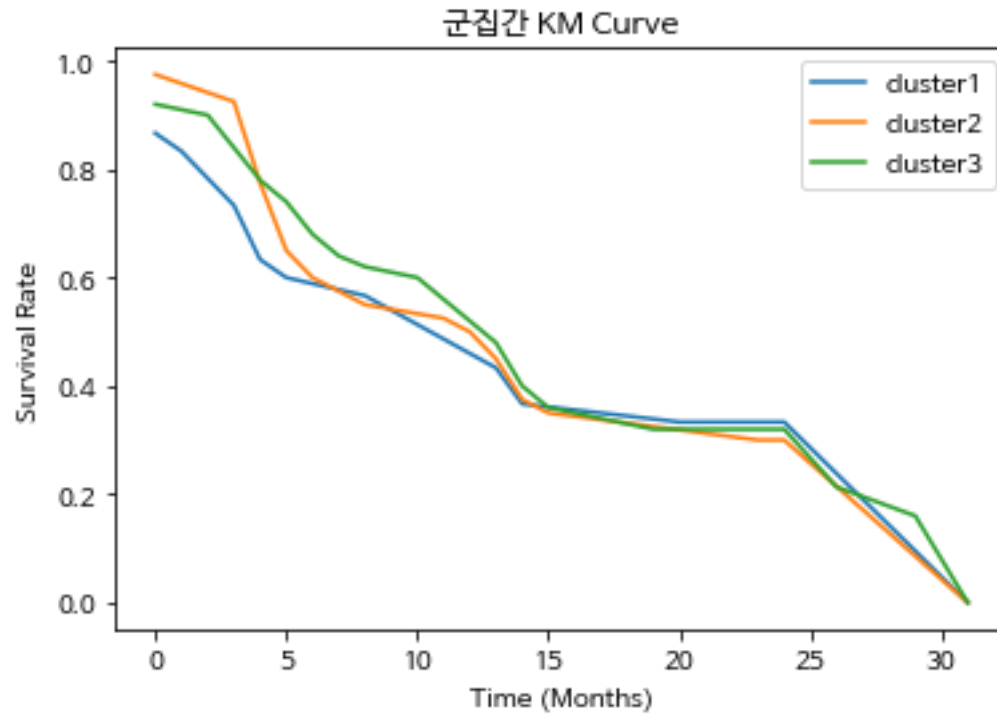


Accelerated Failure  
Time model



## 생존분석 : KM Curve

### Kaplan-Meier Curve



시각적으로는 군집별 생존율이 차이가 있어보이나, 실제로 유의미한 차이가 있을까?



## 생존분석 : KM Curve



Kaplan-Meier Curve

군집간 KM Curve

집단간 생존율 차이에 대한 검정

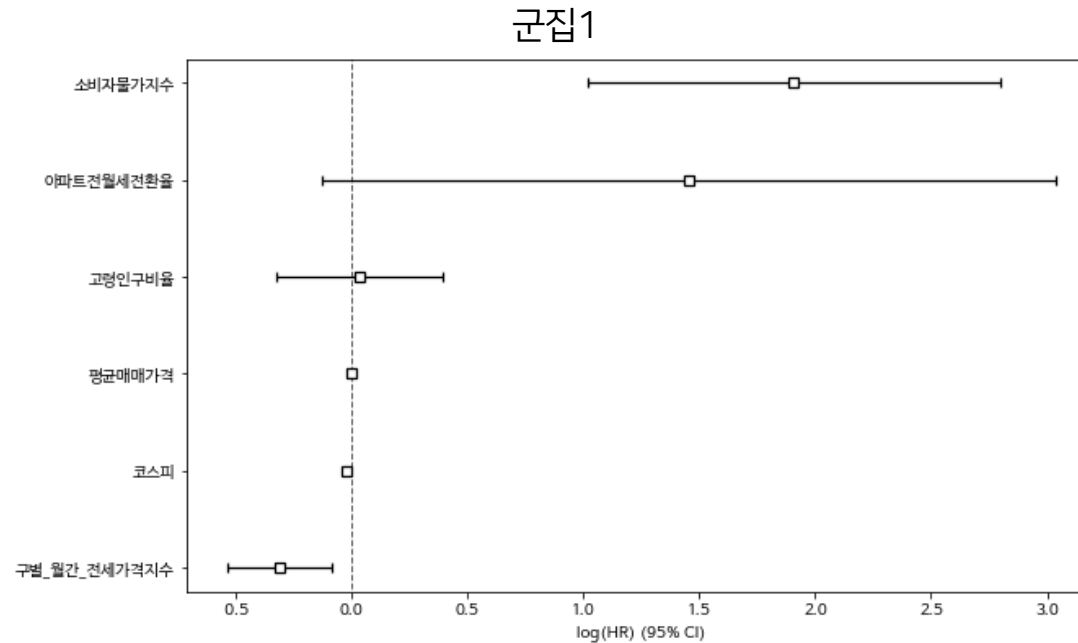
군집1 vs. 군집2	군집2 vs. 군집3	군집1 vs. 군집3
<p>P-value</p> <p>0.92</p>	<p>P-value</p> <p>0.96</p>	<p>P-value</p> <p>0.91</p>

시각적으로는 군집별 생존율이 차이가 있어보이나 실제로 유의미한 차이가 있을까?  
로그순위법 검정 결과, 3개 군집 모두 생존율에 유의미한 차이가 없음을 확인!



## 생존분석 : Cox Proportional Hazard model

### Cox Proportional Hazard model



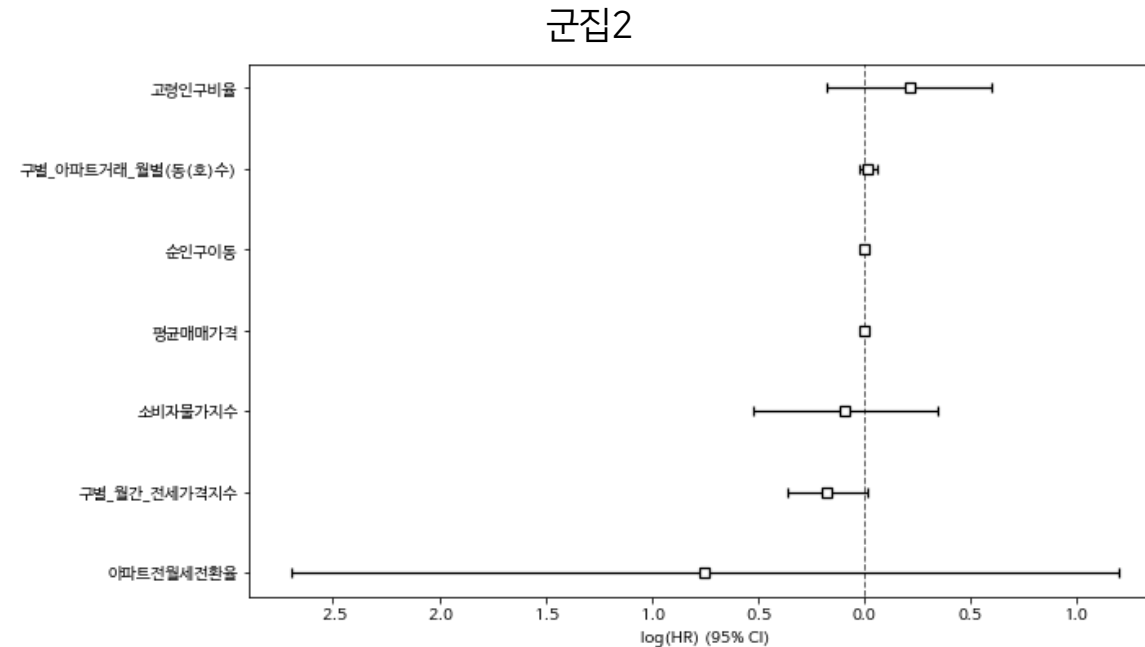
소비자물가지수가 한 단위 **증가** → 역전세 발생 확률 **증가**

구별 월간 전세가격지수가 한 단위 **증가** → 역전세 발생 확률 **감소**



## 생존분석 : Cox Proportional Hazard model

### Cox Proportional Hazard model

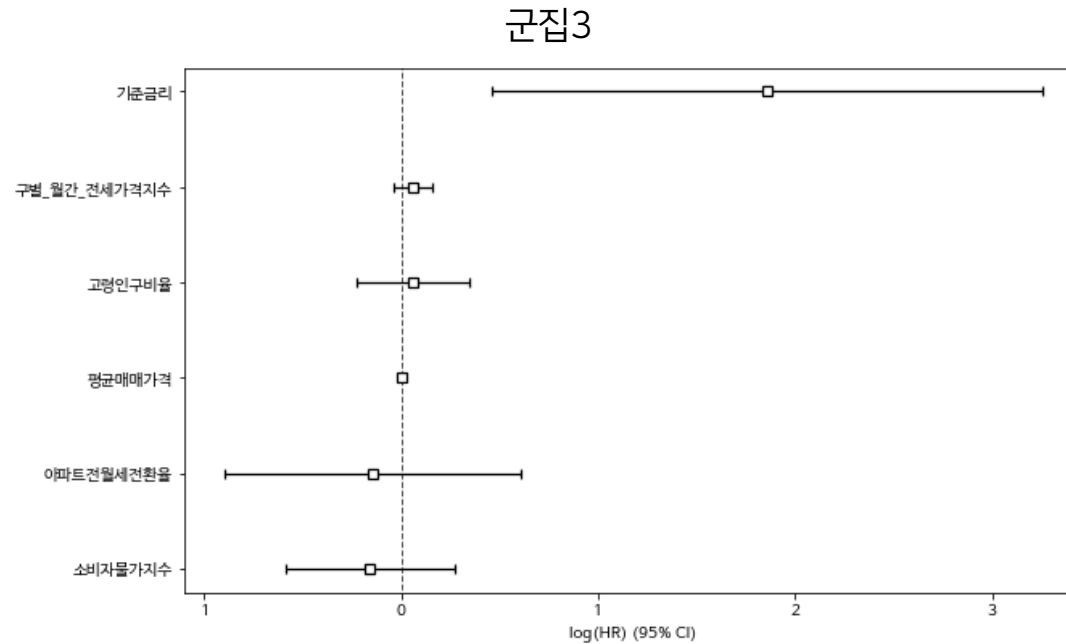


log(위험비) 값들이 모두 0에 근접하거나, 신뢰구간이 0을 포함하고 있으므로  
역전세 발생 확률을 증가/감소시킨다고 보기 어려움!



## 생존분석 : Cox Proportional Hazard model

### Cox Proportional Hazard model



기준금리가 한 단위 **증가** → 역전세 발생 확률 **증가**



## 생존분석 : Accelerated Failure Time model

발생 시간 가속/감속 →

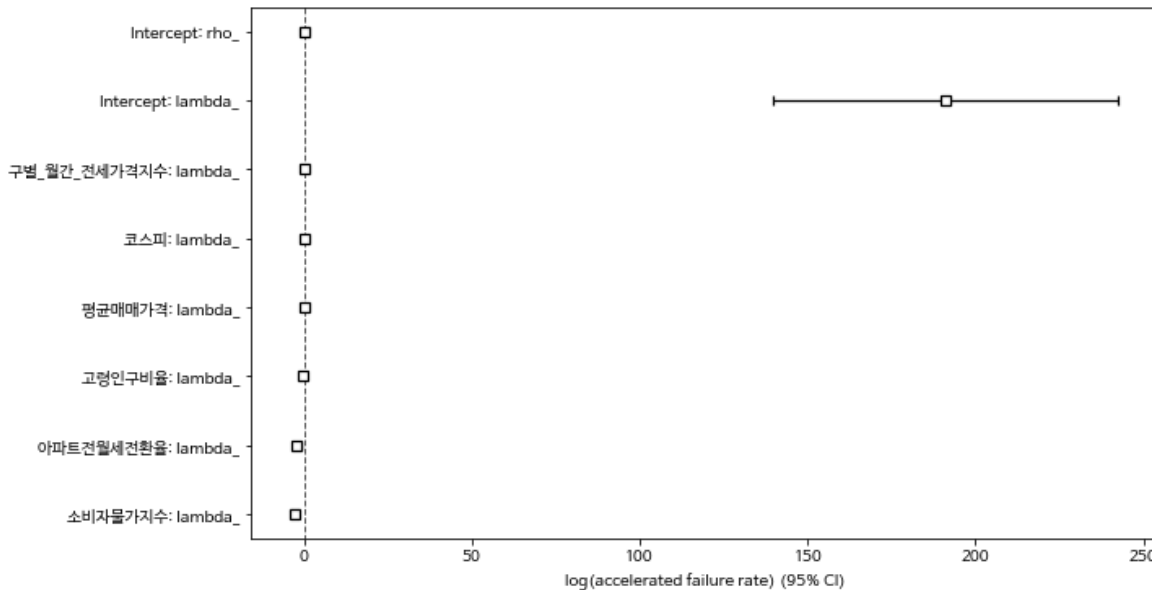
사건이 발생할 때까지의 시간을 가속/감속

(즉, 사건발생확률이 증가/감소)

### Accelerated Failure Time model

1 이상 : 감속 / 1 이하 : 가속

군집1



소비자물가지수, 아파트전월세전환율이 한 단위 **증가** → 역전세 발생 시간 **가속**

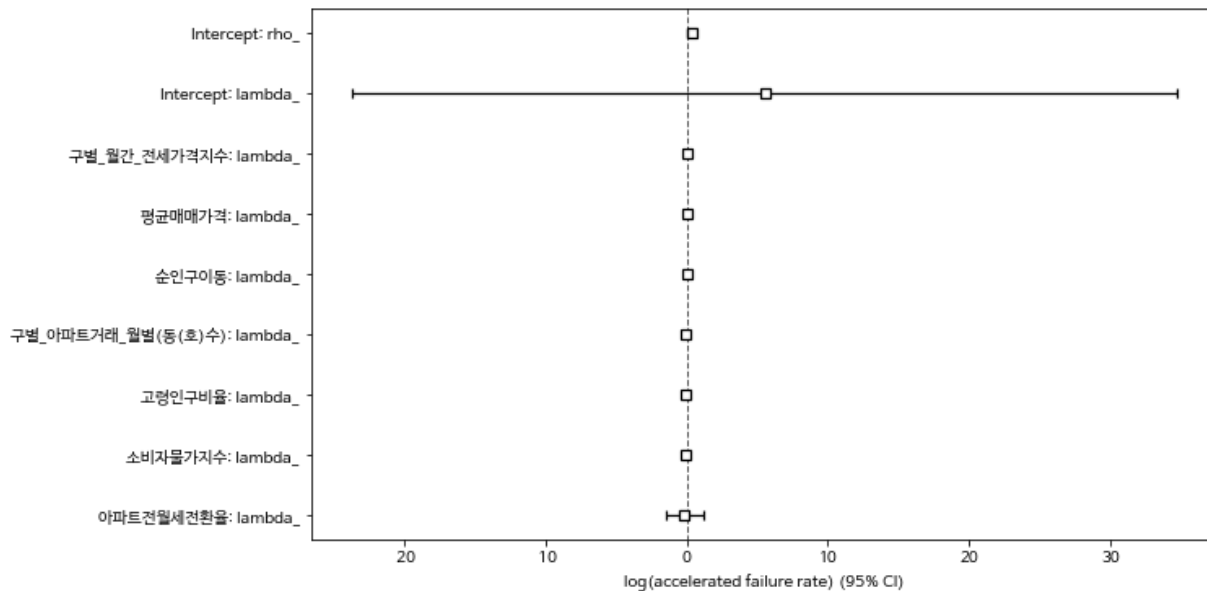
구별 월간 전세가격지수가 한 단위 **증가** → 역전세 발생 시간 **감속**



## 생존분석 : Accelerated Failure Time model

### Accelerated Failure Time model

군집2



변수	Accelerated failure rate
고령인구비율	0.92
아파트거래량	0.98
전세가격지수	1.09
소비자물가지수	0.92
순인구이동	1.00
아파트전월세전환율	0.88

변수별 Accelerated failure rate가 모두 1에 근접하므로,  
역전세 발생을 가속/감속시킨다고 어려움!

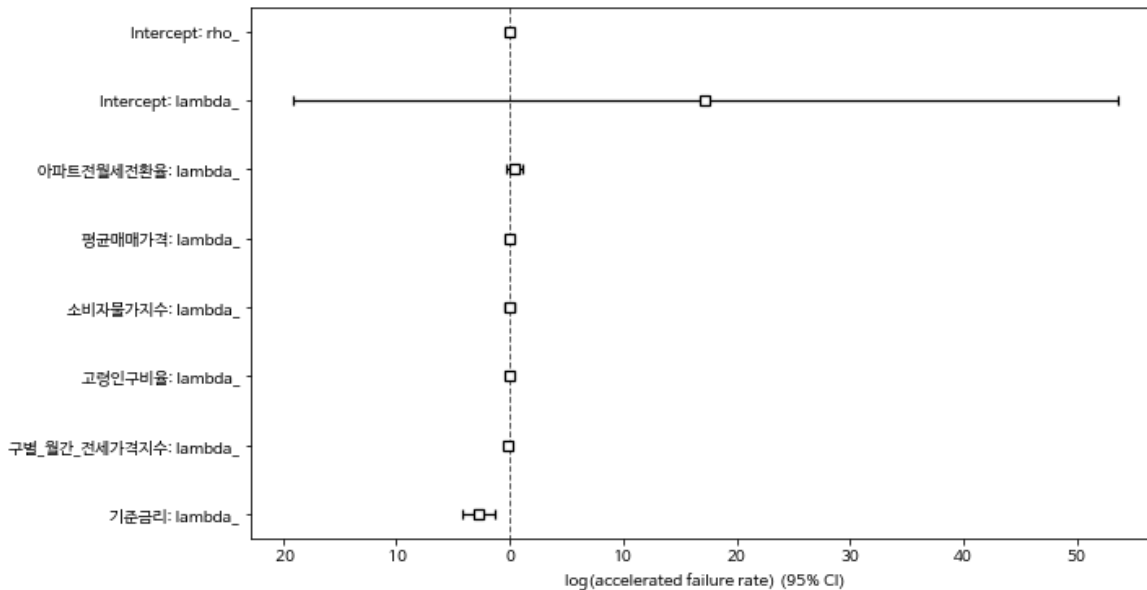




## 생존분석 : Accelerated Failure Time model

### Accelerated Failure Time model

군집3



변수	Accelerated failure rate
고령인구비율	0.94
전세가격지수	0.91
기준금리	0.07
소비자물가지수	0.98
아파트전월세전환율	1.49
평균매매가격	1.00

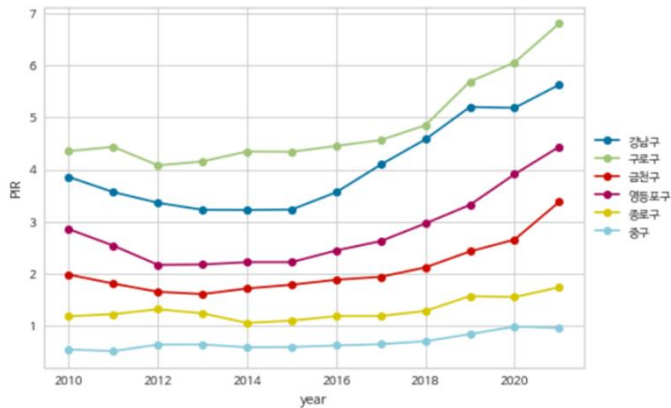
기준금리가 한 단위 **증가** → 역전세 발생 시간 **가속**



## 생존분석 : 결과

### Cox PH + AFT 결과 종합

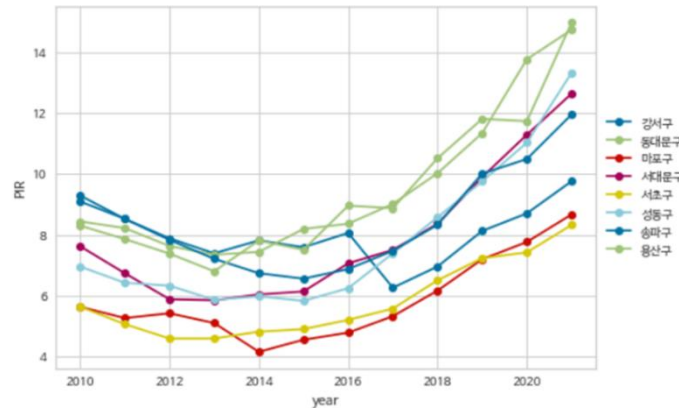
PIR 완만후 상승



소비자물가지수 **증가** → 역전세 확률 **증가**

전세가격지수 **증가** → 역전세 확률 **감소**

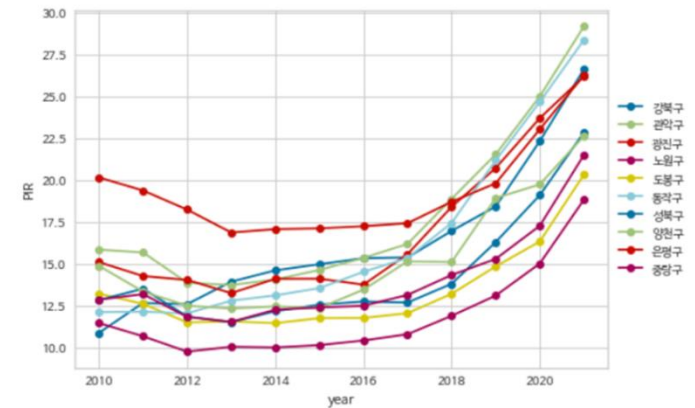
PIR 하락후 급상승



역전세 확률에 영향을 미치는

유의미한 변수 X

PIR 완만후 급상승



기준금리 **증가** → 역전세 확률 **증가**



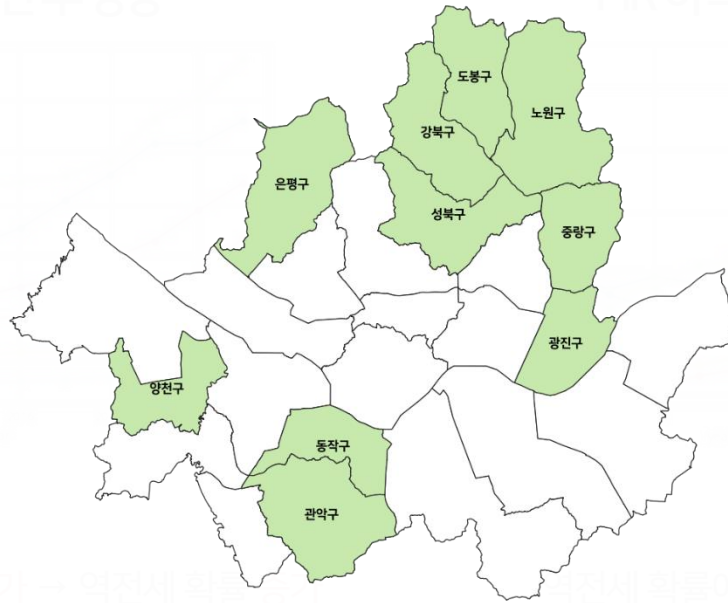
정책 제언 및 한계와 의의



## 정책 제언 및 한계와의의

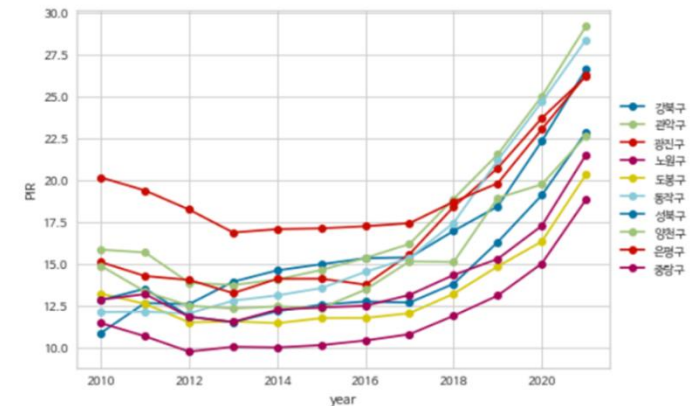
### 정책 제언

PIR 원만후 상승



PIR 하락후 급상승

PIR 원만후 급상승



소비자물가지수 증가 → 역전세 확률 증가

전세가격지수 증가 → 주거부담이 큰 군집에 대한 정책 제언 고려!

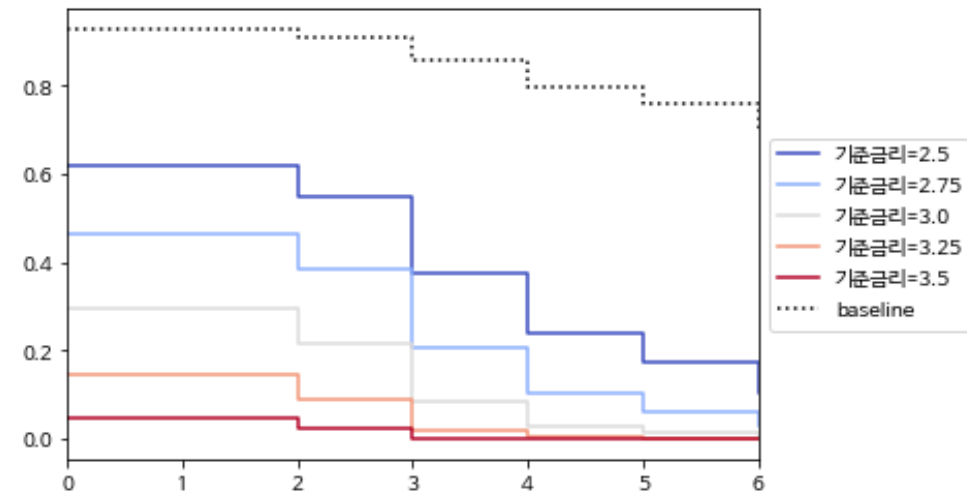
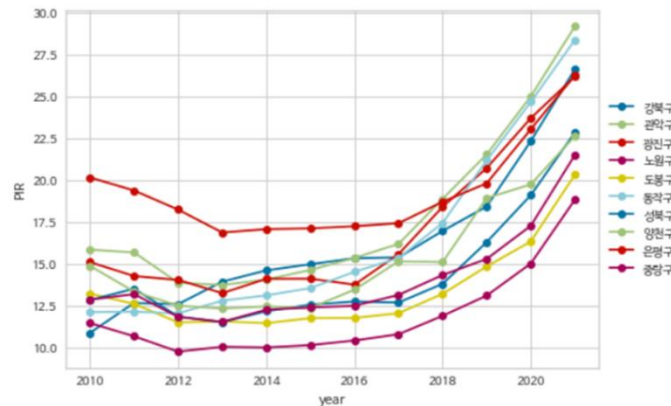
기존금리 증가 → 역전세 확률 증가



## 정책 제언 및 한계와 의의

### 정책 제언

PIR 완만 후 급상승



기준금리가 0.25씩 증가함에 따라

해당 군집에서 단기 생존율이 감소할 것임을 확인!

`plot_partial_effects_on_outcome()` 함수 사용!



## 정책 제언 및 한계와의의

### 정책 제언

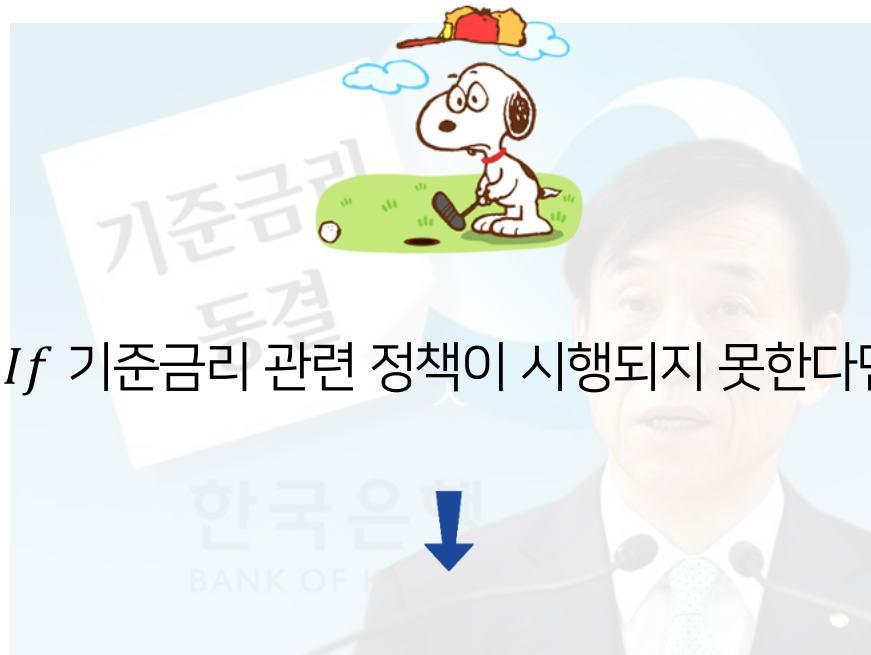


결과적으로, 역전세 위험 완화를 위한 **기준금리 인상폭 감소 또는 동결 정책** 촉구



## 정책 제언 및 한계와의의

### 정책 제언



*If* 기준금리 관련 정책이 시행되지 못한다면...

역전세 피해 구제를 목표로, 피해 대응 정책 수립 필요!

역전세 위험 완화를 위한 기준금리 인상폭 감소 또는 동결 정책 촉구



## 정책 제언 및 한계와의의

### 정책 제언



#### 전세금반환보증보험 확대

임대인의 전세금반환 능력이 없는 경우,  
보증기관의 대위변제 보험상품 개발



#### 한시적 이율인하 대출상품 개발

임차인의 소재지 확인 시,  
낮은 이율의 대출상품 제안





## 정책 제언 및 한계와 의의

### 한계와 의의

#### 한계

제한적인 변수 수집  
(서울시 구별 월별 자료의 부재)

거시경제 위주 분석 결과 도출

#### 의의

생존분석을 통한 역전세 문제 파악

통념에 부합하는 결과 도출  
(기존금리 상승 ↔ 역전세 확률 증가)

클러스터링 + 분석 결과 토대

금리 관련 정책 제언