
“DFM”을 통한 주택매매가격지수 분석 및 예측

2023020369 장희중

>> Table of contents

- 1 분석 목적**
- 2 수집 데이터**
- 3 분석 과정**
- 4 결론 및 한계점**

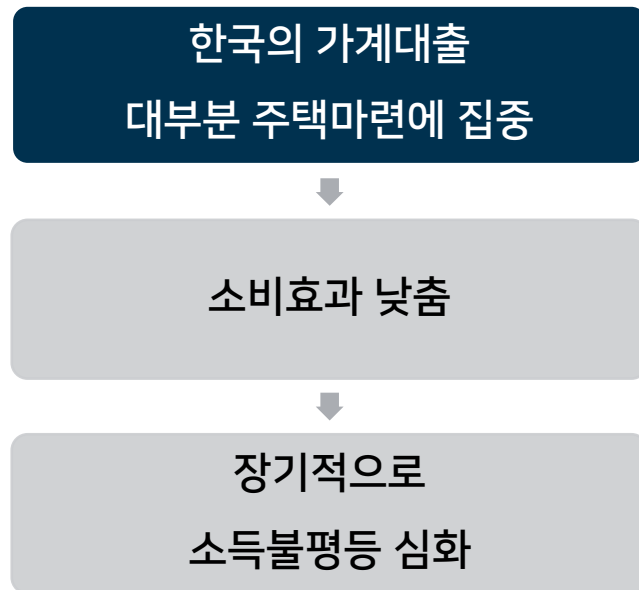
1



분석 목적

Part 1 >> 분석 목적

1. 주택매매가격의 변화 요인 파악



[그림 1] 연도별 신규 가계부채 용도

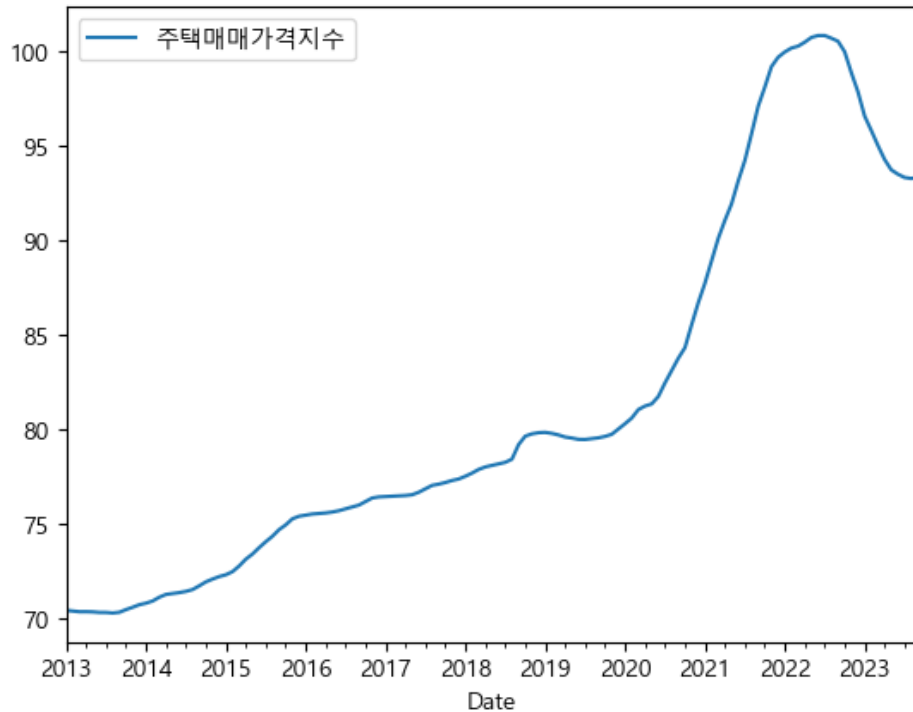


주: 총 표본 5,000가구에 대해 한 가구당 단일 용도만 답변하도록 설계된 설문 결과
자료: 한국노동패널, 저자계산

➡ 주택매매자 및 정부에게 주택매매가격의 변동요인을 파악하는 것은 중요

Part 1 >> 분석 목적

2. 주택매매가격의 동향 예측



주택매매가격 동향 예측



주택매매자의
매매시기 및 대출에
도움을 줌

정부의 정책 방향
결정에 도움

2



수집 데이터

Part 2 >> 수집 데이터 (2013-01 ~ 2023-09)

출처	수집한 데이터
ECOS	1. 주택매매가격지수 (전국 총 지수) 2. 고용률 (경제활동인구) 3. 주택담보대출금리 (예금은행 대출금리) 4. 건설업_업황전망, 건설업_매출전망, 건설업_자금사전전망 5. 소비자 물가지수, 생산자 물가지수 6. 미분양 주택 현황 7. M2 (통화량) 8. 국고채(3년, 10년, 30년), 회사채(3년) 9. KOSPI 증가, KOSDAQ 증가 10. 국채 거래량, 회사채 거래량 11. GDP 12. 주택전세가격지수 13. 환율
KOSIS	주택 거래량 (아파트)

→ 총 27개 데이터 수집

Part 2 >> 수집 데이터

	Date	기준 금리	주택건설 인허가실 적	전산업생 산지수	주택담 보대출 금리	건설 업_업 황전 망	건설 업_매 출전 망	건설업 _자금 사전전 망	소비자물 가지수	생산자물 가지수	...	국고 채(30 년)	회사 채(3 년)	KOSPI_ 증가	KOSDAQ_ 증가	국채 거래 량	회사채 거래량	GDP	주택전 세가격 지수	환율	주택매매 가격지수
0	2013-01-01 00:00:00	2.75	23952	62.2697	4.55	56	68	62	92.728	104.96	...	3.31	3.20	1961.94	503.67	121500624	573916	0.9	82.7	1066.54	70.462
1	2013-02-01 00:00:00	2.75	45543	61.6258	4.51	52	76	77	93.038	105.68	...	3.30	3.06	2026.49	535.87	99040466	422222	0.9	82.9	1086.06	70.421
2	2013-03-01 00:00:00	2.75	75594	76.0634	4.46	58	68	72	92.952	105.26	...	3.17	2.95	2004.89	555.02	104760077	481596	0.9	83.2	1103.81	70.389
3	2013-04-01 00:00:00	2.75	106509	80.5564	4.40	61	71	74	92.823	104.93	...	3.15	2.91	1963.95	563.87	110564496	335381	1.2	83.6	1121.10	70.391
4	2013-05-01 00:00:00	2.50	142055	81.3058	4.33	58	74	70	92.823	104.57	...	3.17	2.96	2001.05	577.87	117919540	531677	1.2	83.8	1112.10	70.381

3



분석 과정

Part 3 >> 분석 과정



Granger Causality Test

Part 3 >> Granger Causality Test

- 수집한 데이터들에 대해서 주택매매가격지수(y) 예측에 변수들(x)의 과거 값이 유의한지 검정

H0: x로부터 y로의 인과방향이 존재하지 않음

H1: x로부터 y로의 인과방향이 존재한다고 할 수 있음

=> p-value가 0.05보다 작으면 x로부터 y로의 인과방향이 존재한다고 판단.

변수명(x)	P-value	변수명(x)	P-value
기준금리	0.0	회사채(3년)	0.0
주택건설인허가실적	0.035	KOSPI_증가	0.0
전산업생산지수	0.04	KOSDAQ_증가	0.0
주택담보대출금리	0.0	국채 거래량	0.0498
건설업_업황전망	0.0001	회사채 거래량	0.4759
건설업_매출전망	0.0024	GDP	0.024
건설업_자금사전전망	0.0	주택전세가격지수	0.0001
소비자물가지수	0.0	환율	0.0
생산자물가지수	0.0	주택 거래량	0.0
미분양주택현황	0.0	고용률	0.0006
국고채(3,10,30년)	0.0	M2	0.0007

회사채 거래량 변수만 p-value가 0.05보다 크고
국채 거래량 또한 0.05로 비교적 큰 p-value가짐

→ 국채 거래량, 회사채 거래량 변수 제거

Principal Component Analysis (PCA)

Part 3 >> PCA

수집한 변수들을 주택매매가격지수, 기준금리를 제외하고 **총 8가지의 그룹**으로 분리

주택건설인허가실적
건설업_업황전망
건설업_매출전망
건설업_자금사전전망

① 건설업

주택담보대출금리
미분양주택현황

② 부동산 시장

국고채(3년)
국고채(10년)
국고채(30년)
회사채(3년)

③ 채권금리

KOSPI 증가
KOSDAQ 증가

④ 주식시장

소비자물가지수
생산자물가지수

⑤ 물가

환율
전산업생산지수
고용률
M2
GDP

⑥ 경제

⑦ 주택전세가격지수

⑧ 주택거래량

Part 3 >> PCA: Explained Ratio

그룹별로 PCA를 적용하여 차원 축소

그룹별로 principal component의 개수에 따른 설명 비율은 다음과 같음

Explained ratio		
Group	Principal Component	
	PC1	PC2
건설업	0.57	0.26
부동산 시장	0.78	0.22
채권 금리	0.93	0.05
주식 시장	0.94	0.06
물가	0.86	0.14
경제	0.45	0.18

➡ 6개 그룹 모두 첫번째 principal component만 사용하기로 결정

Part 3 >> PCA

- PCA를 통해 차원 축소된 데이터를 재정의

기준금리는 외생변수로 설정

	건설업_pc1	부동산 시장_pc1	채권금리_pc1	주식시장_pc1	물가_pc1	경제_pc1	주택전세가격지수	주택 거래량	주택매매가격지수
Date									
2013-01-01	-0.741866	0.778400	0.613032	-0.440554	-0.235296	-0.730172	0.000000	0.000000	0.004551
2013-02-01	-0.451541	0.742935	0.576023	-0.368113	-0.200063	-0.710044	0.009709	0.068153	0.003208
2013-03-01	-0.523430	0.692743	0.484398	-0.351124	-0.216168	-0.553820	0.024272	0.175387	0.002161
2013-04-01	-0.368974	0.665479	0.448577	-0.356738	-0.231702	-0.436456	0.043689	0.281171	0.002226
2013-05-01	-0.422549	0.601515	0.481964	-0.321312	-0.242049	-0.410529	0.053398	0.374761	0.001899

Dynamic Factor Model (DFM)

Part 3 >> DFM

- Dynamic Factor Model

소수의 잠재 요인이 더 많은 수의 관측된 시계열의 공통적인 역학을 설명한다고 가정

$$X_t = \Lambda F_t + A x_t + u_t$$

$$F_t = B_1 F_{t-1} + B_2 F_{t-2} + \eta_t$$

$$u_t = C_1 u_{t-1} + C_2 u_{t-2} + \epsilon_t$$

X_t : Endogenous var (기준금리 제외 모든 변수)

F_t : Latent Factor

x_t : Exogenous var (기준금리)

u_t : 개별 요인

```
endog = train_total
exog = X_train_sc[['기준금리']]

# DFM 모델 정의
n_factors = 1
dfm = DynamicFactor(endog=endog, exog = exog, k_factors=n_factors, factor_order=2, enforce_stationarity=True, error_order=2)
dfm_results = dfm.fit(maxiter=500)
```

AIC를 기준으로 factor 개수와 factor lag를 튜닝한 결과,
n_factors = 1, factor order=2로 결정
잔차의 자기상관문제를 해결하기 위해 **error order=2**로 설정

Part 3 >> DFM

- 분석 결과

1. Λ : Coefficient of Latent Factor

변수	Coef of Latent factor F	P-value
건설업	1.62	0.283
부동산 시장	-2.31	0.0
채권금리	-1.34	0.394
주식시장	2.62	0.126
물가	2.13	0.0
경제	2.9	0.011
주택 전세가격지수	0.23	0.08
주택 거래량	0.50	0.69
주택매매가격지수	0.16	0.031

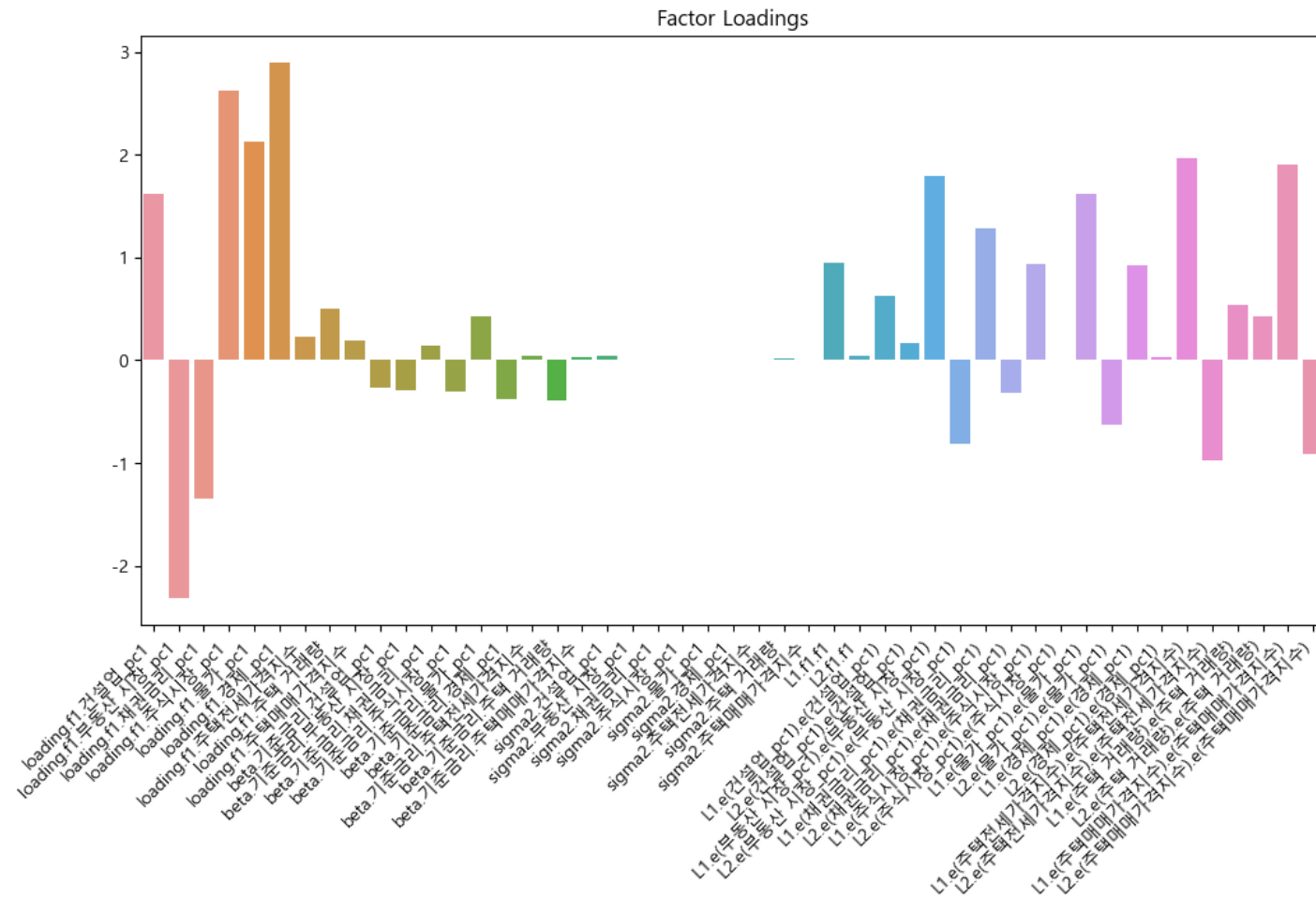
2. B: coefficient of lag of Latent Factor

Lag of Latent factor	Coef of lag.F	P-value
F_{t-1}	0.95	0.0
F_{t-2}	0.04	0.24

$$\begin{aligned}X_t &= \Lambda F_t + A x_t + u_t \\F_t &= B_1 F_{t-1} + B_2 F_{t-2} + \eta_t \\u_t &= C_1 u_{t-1} + C_2 u_{t-2} + \epsilon_t\end{aligned}$$

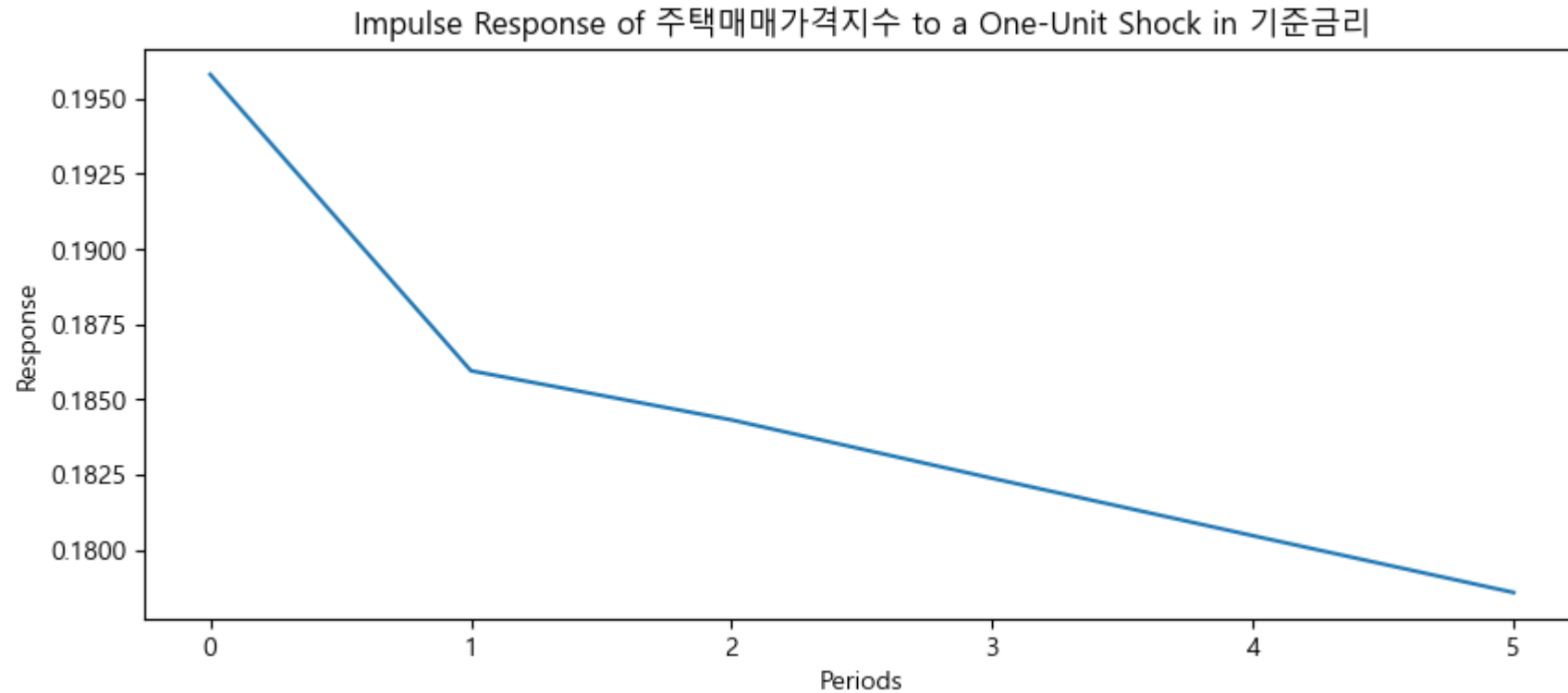
Part 3 >> DFM

- 분석 결과



Part 3 >> DFM

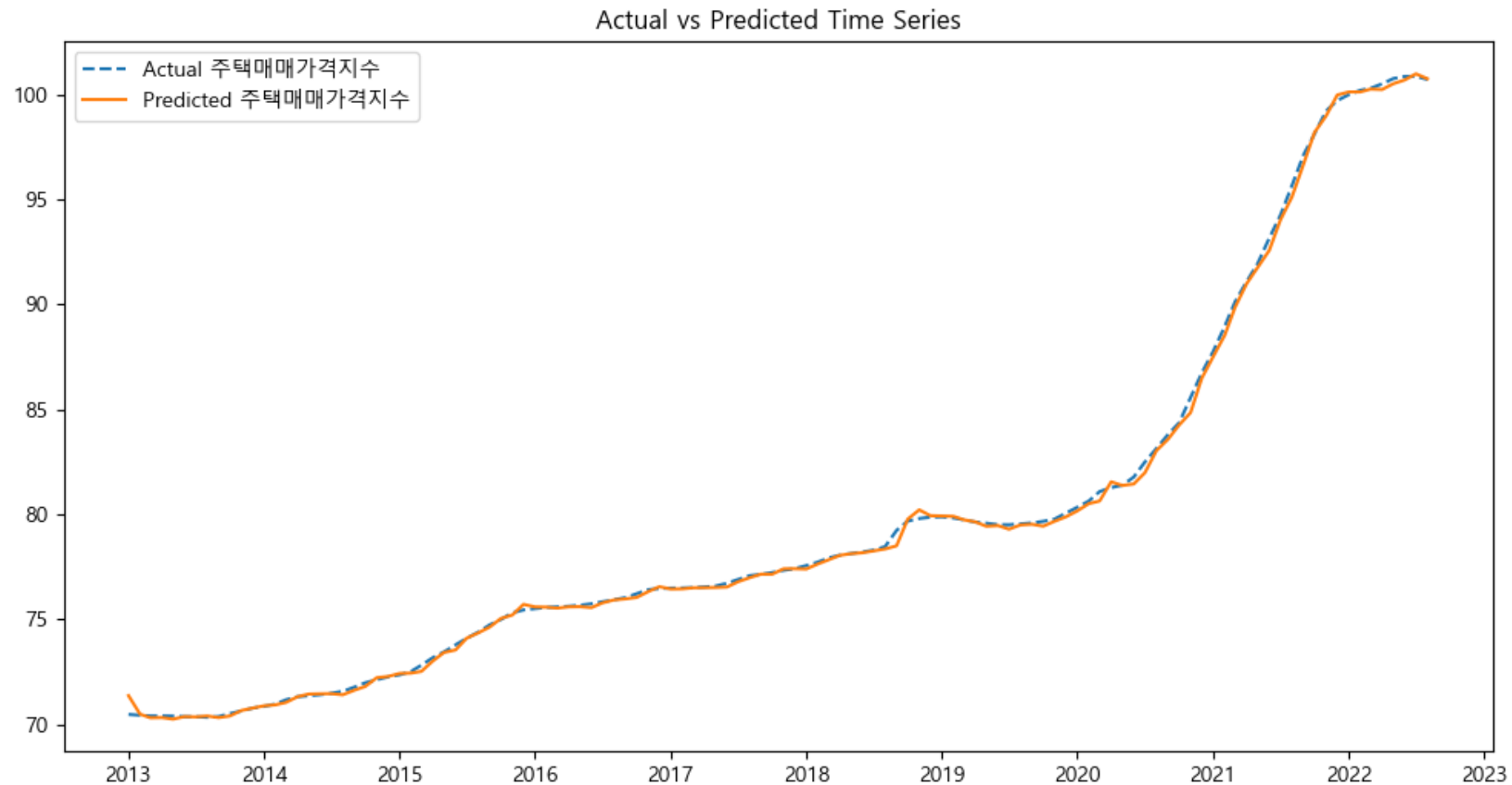
- Impulse Response



→ 기준금리의 one-unit shock은 주택매매가격지수의 하락을 초래

Part 3 >> DFM

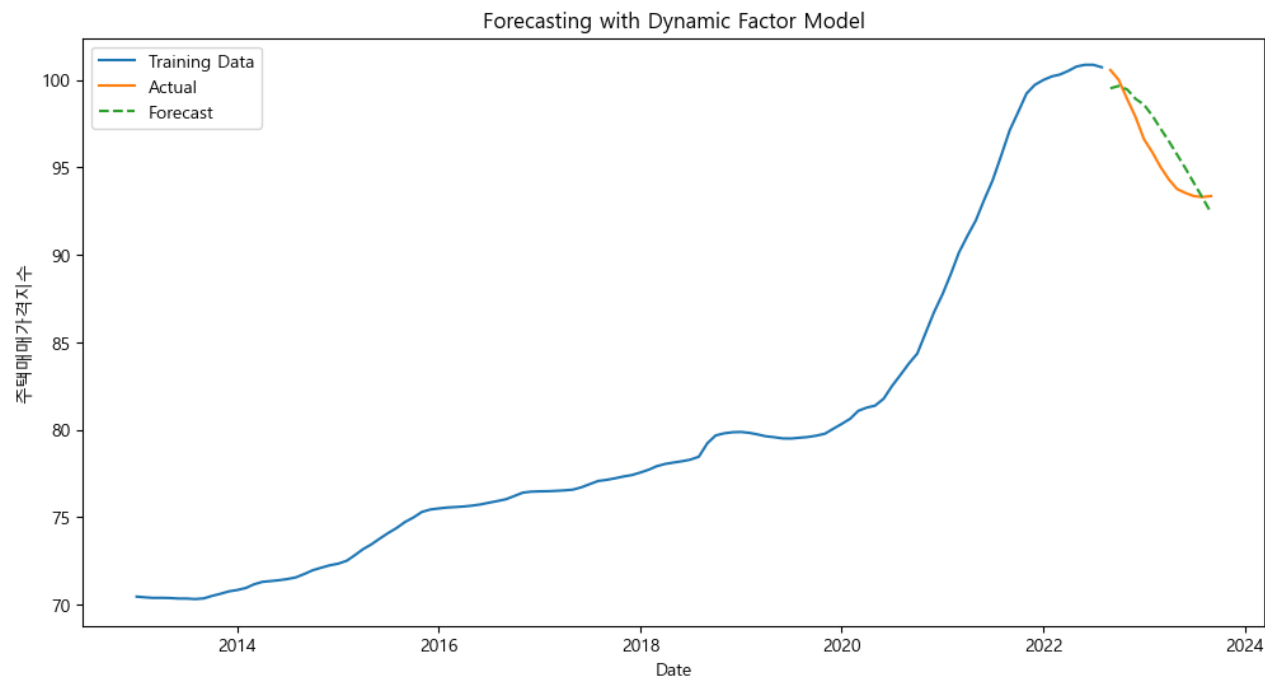
- In-sample에 대한 예측정확도



Train RMSE: 0.215

Part 3 >> DFM

- Out-of-sample에 대한 예측 정확도

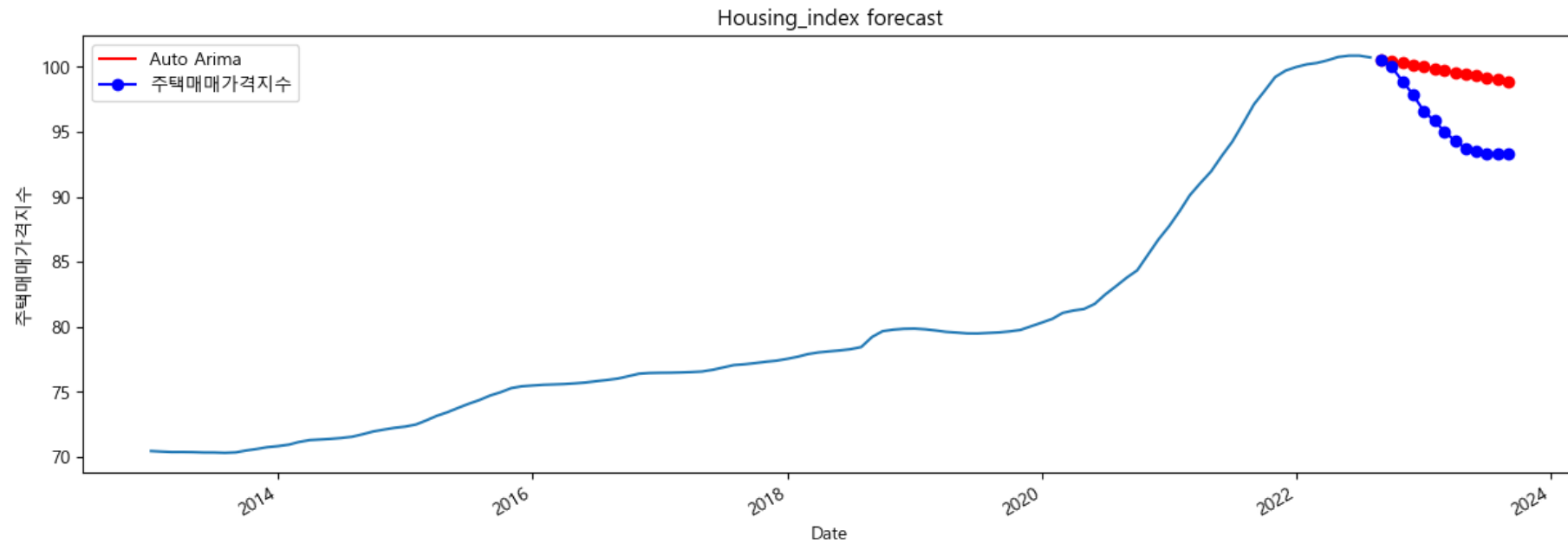


	실제 주택매매가격지수	예측 매매가격지수
2022-09-01	100.568	99.522438
2022-10-01	100.014	99.659512
2022-11-01	98.909	99.457822
2022-12-01	97.894	98.926172
2023-01-01	96.612	98.589708
2023-02-01	95.848	97.939671
2023-03-01	95.054	97.241898
2023-04-01	94.316	96.503412
2023-05-01	93.766	95.730785
2023-06-01	93.536	94.930156
2023-07-01	93.360	94.107240
2023-08-01	93.308	93.267340
2023-09-01	93.365	92.415368

Test RMSE: 1.46

Part 3 >> 모델 비교

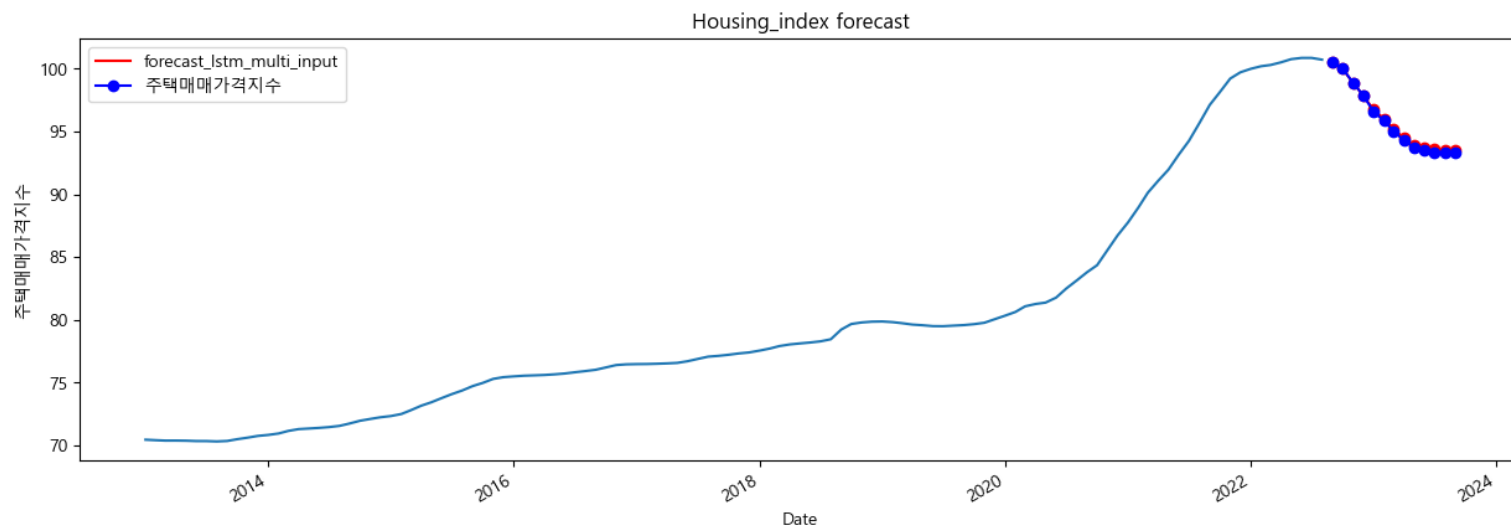
- ARIMA



Test RMSE: 4.369

Part 3 >> 모델 비교

- 단순히 예측정확도를 비교하기 위해 LSTM과 비교



	실제 주택매매가격지수	예측 매매가격지수
Date		
2022-09-01	100.568	100.661636
2022-10-01	100.014	100.146828
2022-11-01	98.909	99.157486
2022-12-01	97.894	98.262238
2023-01-01	96.612	96.553162
2023-02-01	95.848	95.838493
2023-03-01	95.054	95.122681
2023-04-01	94.316	94.330658
2023-05-01	93.766	93.663544
2023-06-01	93.536	93.608498
2023-07-01	93.360	93.248817
2023-08-01	93.308	93.242798
2023-09-01	93.365	93.422150

Test RMSE: 0.15

4



결론 및 한계점

Part 4 >> 결론 및 한계점

- 결론

1. 주택매매가격은 건설업, 부동산시장, 물가 등 여러 복합적 요인으로 결정됨.

DFM 모형은 이 복합적 요인들의 효과를 계수추정치들을 통해 잘 보여줌.

2. 정확한 집값이 아닌 주택매매가격지수의 예측이기 때문에

DFM 모형의 예측정확도도 주택매매가격지수의 미래 동향을 보기에 좋은 편이라 생각됨

다만 아주 정확한 예측을 목적으로 한다면 LSTM과 같은 딥러닝 모형을 사용하는 것이 좋음

→ 주택매매가격지수에 대한 분석 및 예측을 위해 DFM 모형을 활용하는 것은 합리적

- 한계점

더 다양하고 적절한 데이터의 수집으로 좀 더 합리적인 그룹화를 통한 PCA 수행 후 DFM 모형을 적용한다면 주택매매가격의 복합적 요인을 더 구체적이고 정확하게 분석할 수 있을 것

Thank you