
머신 러닝 알고리즘을 이용한 부동산 가격 결정 요인 분석 및 판매 가격 예측

Analysis of Real Estate Price Determinants
and Prediction of Sales Price Using
Machine Learning Algorithm

IT경영학과 2014314005 김나은
지도교수 강지훈

2020. 5. 26

목차

CONTENTS



01 연구 개요

- 1 □
- 예측 부동산 가격($t+k$) = $f(X, Y, \text{부동산 종류}, g(\text{경제지표}, t))$

- 2 탐색적분석테마
- 금리와 부동산 가격 고찰
 - 환율과 부동산 가격 고찰

- 3 내가 세운 가설
- 부동산 구매 요인은 두가지로 나뉨(거주 목적, 투자 목적)
 - 거주 목적과 투자 목적의 부동산의 가격은 서로 다를 것이다.
 - 금리 이외의 다른 변수들이 부동산 가격에 영향을 끼침
 - 환율과 부동산 가격이 반비례 \propto

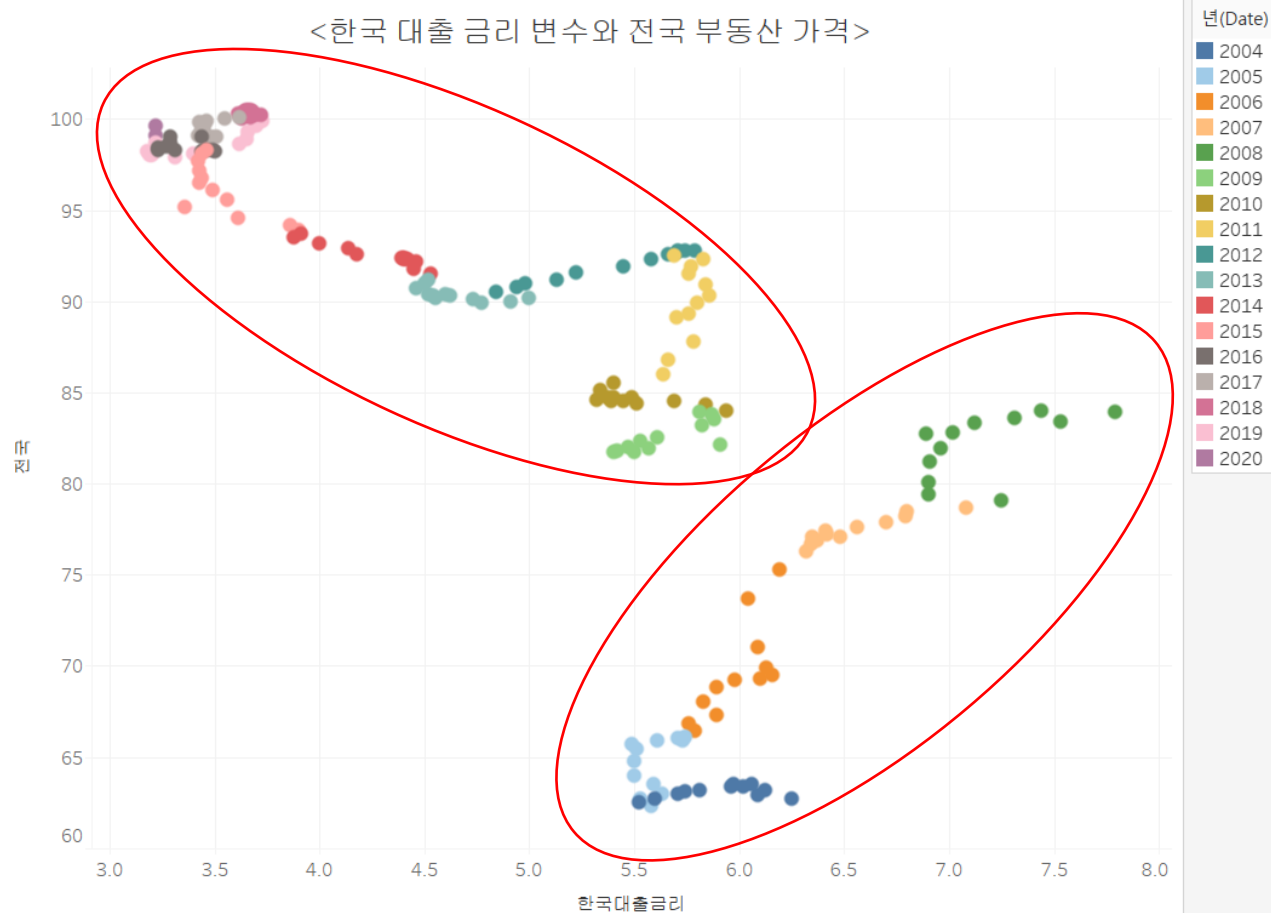
머신러닝 기반 부
동산 예측 모델링

02 연구 배경

부동산 판매가격과 한국 대출 금리 관계를 거시적 관점으로 분석

가설 : 낮은 금리는 부동산 판매 가격을 상승시킨다.

〈그림 1〉 한국 대출 금리 변수 데이터와 전국 부동산 판매 가격 데이터 scatter plot



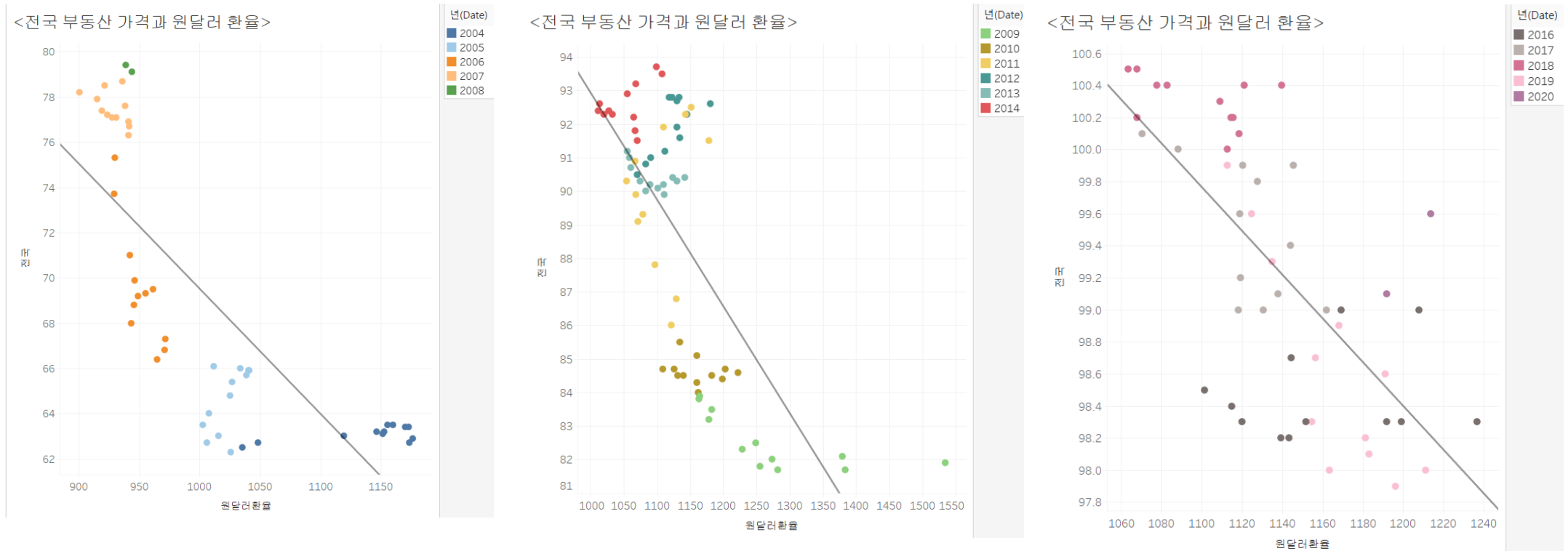
- 2004-01-01부터 2020-02-01기간의 한국 대출 금리와 전국 부동산 가격 평균 데이터를 scatter plot으로 나타냈다
- 2008년 글로벌 금융위기 전후로 금리와 전국 부동산 가격은 멀티모달(Multi-Modal)의 형태로 나타났다.
- 낮은 금리는 부동산 판매 가격을 상승시킨다는 일반적인 통념과 다른 형태로 나타난다.

02 연구 배경

부동산 판매가격과 원달러환율과의 관계를 거시적 관점으로 분석

가설 : 원 달러 환율이 오르면 부동산 가격이 하락한다.

<그림 2> 구간(연도) 별 원달러환율 데이터와 전국 부동산 판매 가격 데이터 scatter plot



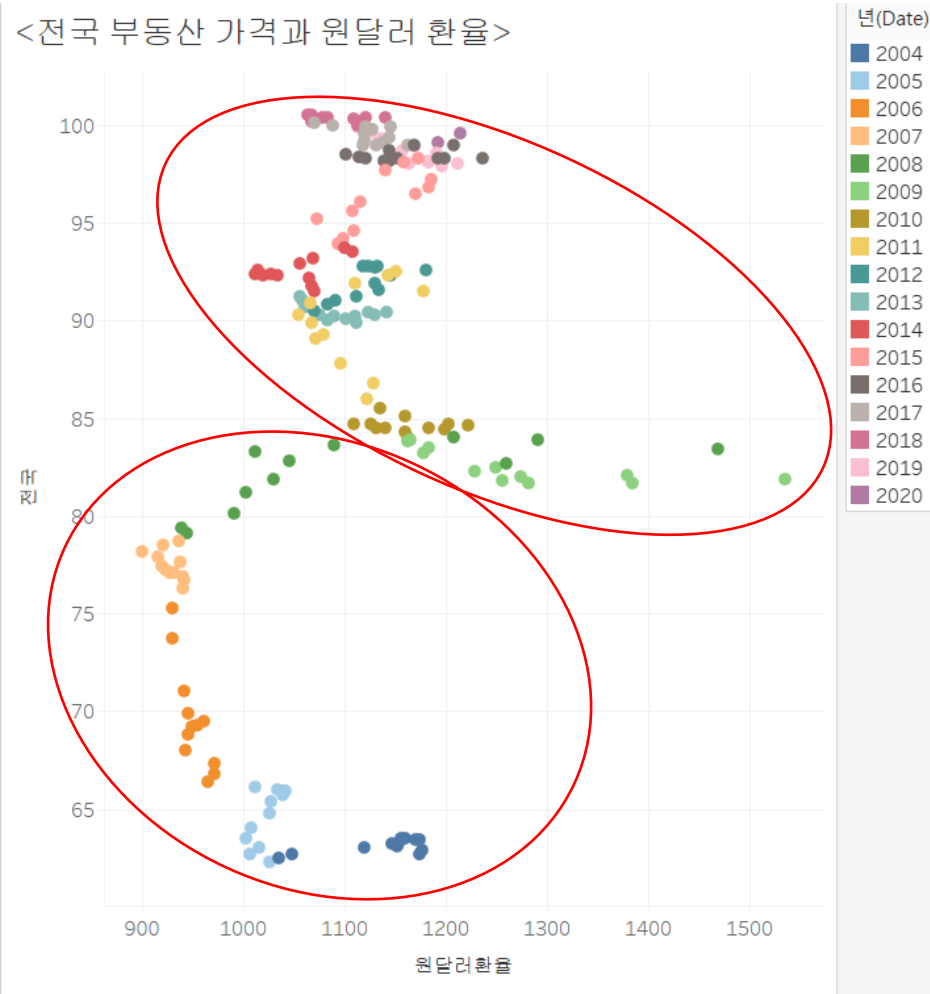
- 2004-01-01부터 2020-02-01기간의 원달러환율 데이터와 전국 부동산 가격 평균 데이터를 구간(연도)별로 나누어 scatter plot으로 나타냈다.
- 원 달러 환율이 오르면 부동산 가격이 하락한다는 가설이 맞는 것 같아 보인다 ㅋ.

02 연구 배경

부동산 판매가격과 원달러환율과의 관계를 거시적 관점으로 분석

가설 : 원 달러 환율이 오르면 부동산 가격이 하락한다.

〈그림 3〉 원달러환율 데이터와 전국 부동산 판매 가격 데이터 scatter plot



- 2004-01-01부터 2020-02-01기간의 원달러환율 데이터와 전국 부동산 가격 평균 데이터를 scatter plot으로 나타냈다.
- 원 달러 환율이 오르면 부동산 가격이 하락한다는 가설을 설명하기 힘들다.
- 2008년 글로벌 금융위기 전후로 원달러환율과 전국 부동산 가격은 멀티모달(Multi-Modal)의 형태로 나타났다.

02 연구 배경

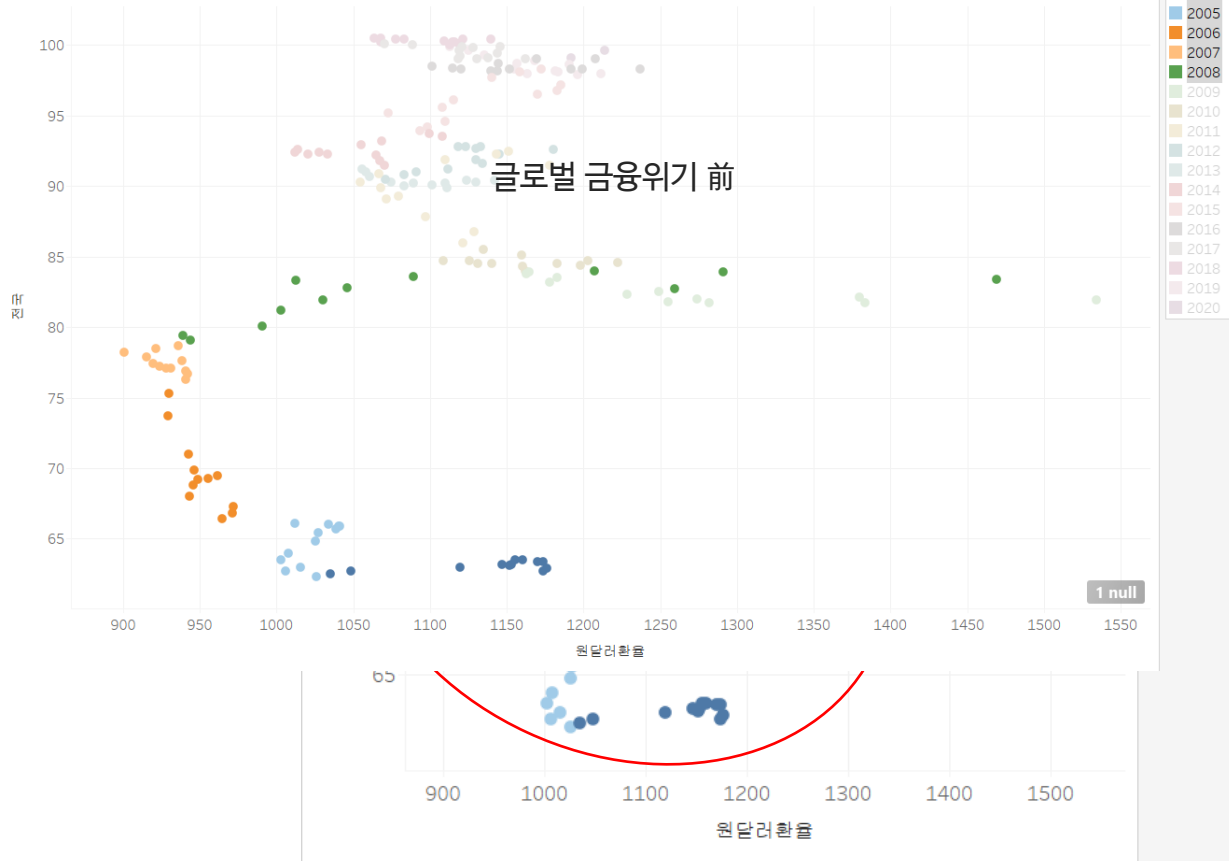
부동산 판매가격과 원달러환율과의 관계를 거시적 관점으로 분석

가설 : 원 달러 환율이 오르면 부동산 가격이 하락한다.

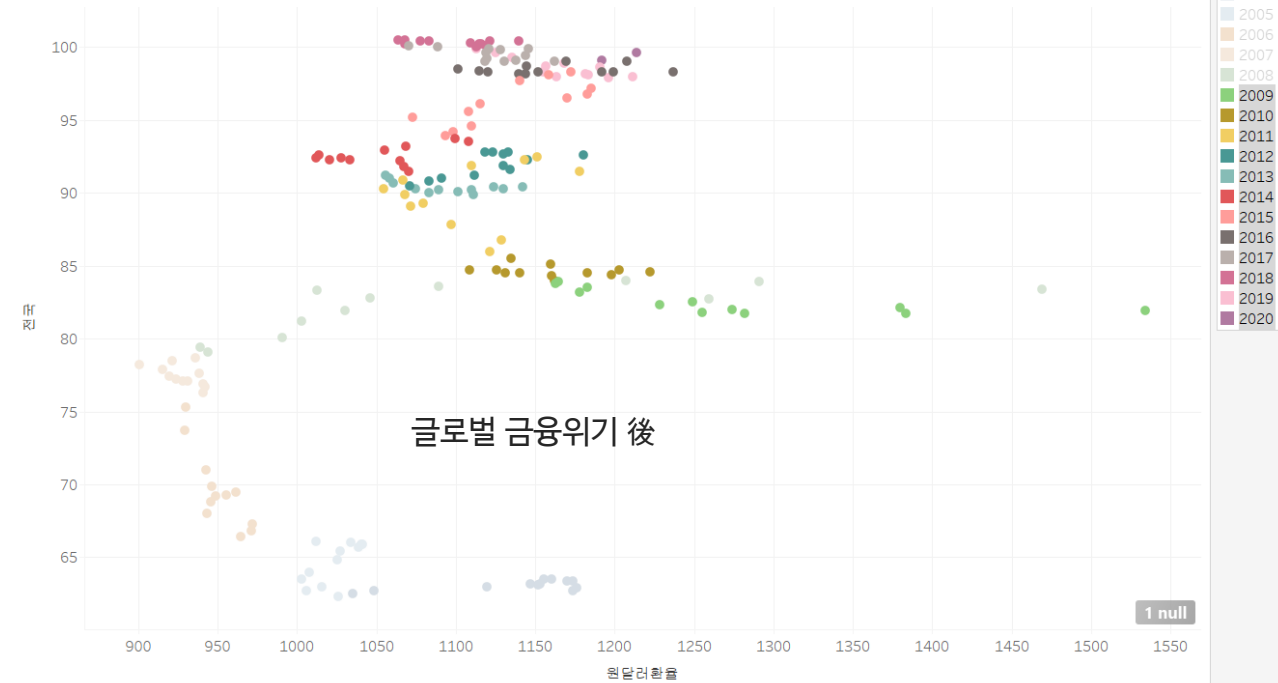
〈그림 3〉 원달러환율 데이터와 전국 부동산 판매 가격 데이터 scatter plot

<전국 부동산 가격과 원달러 환율>

<전국 부동산 가격과 원달러 환율>



<전국 부동산 가격과 원달러 환율>

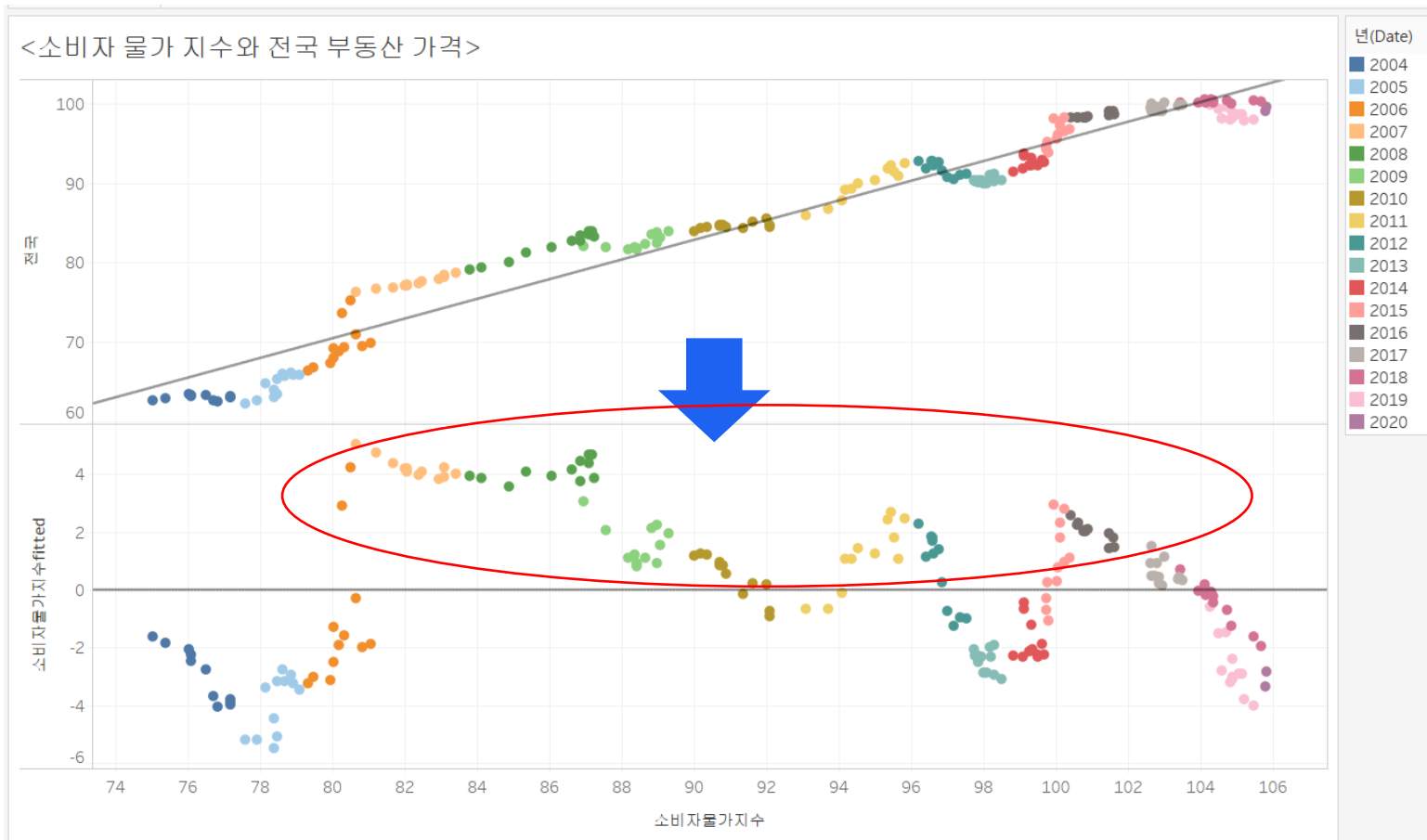


03 현행 분석

탐색적 분석을 통한 소비자물가지수와 전국 부동산 가격

소비자물가지수와 전국 부동산 가격

〈그림 4〉 소비자물가지수데이터와 전국 부동산 판매 가격 데이터 scatter plot



- 2004-01-01부터 2020-02-01기 간의 소비자물가지수 데이터와 전국 부동산 가격 평균 데이터를 scatter plot으로 나타냈다
- 아래 차트는 동일한 데이터로 $y - \hat{y}$ 을 한 Residual이다.
- Residual이 0보다 큰 데이터는 소비자물가지수 대비 부동산 가격이 고평가 된 구간이다.
- 물가는 소비자의 구매력을 의미 하며 Residual이 0보다 큰 데이터는 투자심리가 반영된 부동산 가격이라고 할 수 있다.

03 현행 분석

탐색적 분석을 통한 1인당 국민총소득과 전국 부동산 가격

1인당 국민총소득과 전국 부동산 가격

<그림 5> 1인당 국민 총소득 데이터와 전국 부동산 판매 가격 데이터 scatter plot

<국민총소득과 부동산가격>



- 2004-01-01부터 2019-12-01기간의 연도별 1인당 국민 총 소득 데이터와 2004-01-01부터 2020-02-01 월별 전국 부동산 가격 평균 데이터를 scatter plot으로 나타냈다.
- 아래 차트는 동일한 데이터로 $y - \hat{y}$ 을 한 Residual이다.
- Residual이 0보다 큰 데이터는 1인당 국민총소득 대비 소비자가 과소비하여 부동산을 취득하는 데이터이다.
- Residual이 0보다 큰 데이터는 부동산을 구매할 때 투자심리가 반영되었다고 할 수 있다.

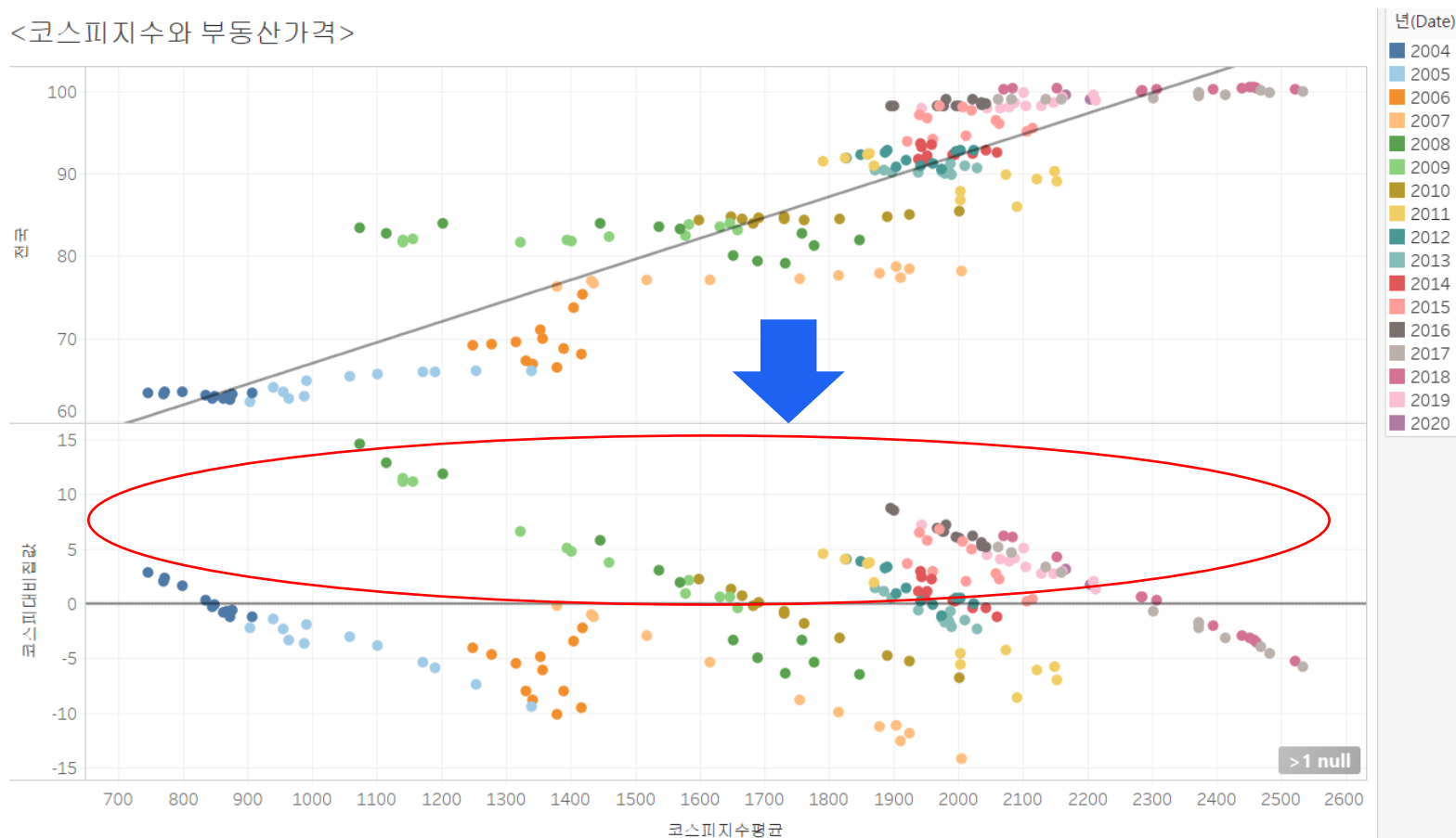
03 현행 분석

탐색적 분석을 통한 코스피지수평균과 전국 부동산 가격

코스피지수평균과 전국 부동산 가격

〈그림 6〉 코스피지수평균 데이터와 전국 부동산 판매 가격 데이터 scatter plot

<코스피지수와 부동산가격>



- 2004-01-01부터 2020-02-01기간의 코스피지수평균 데이터와 전국 부동산 가격 평균 데이터를 scatter plot으로 나타냈다.
- 아래 차트는 동일한 데이터로 $y - \hat{y}$ 을 한 Residual이다.
- Residual이 0보다 큰 데이터는 주식 투자심리에 비해 부동산 투자 심리가 큰 구간이다.
- Residual이 0보다 큰 데이터는 부동산을 구매할 때 투자심리가 반영되었다고 할 수 있다.

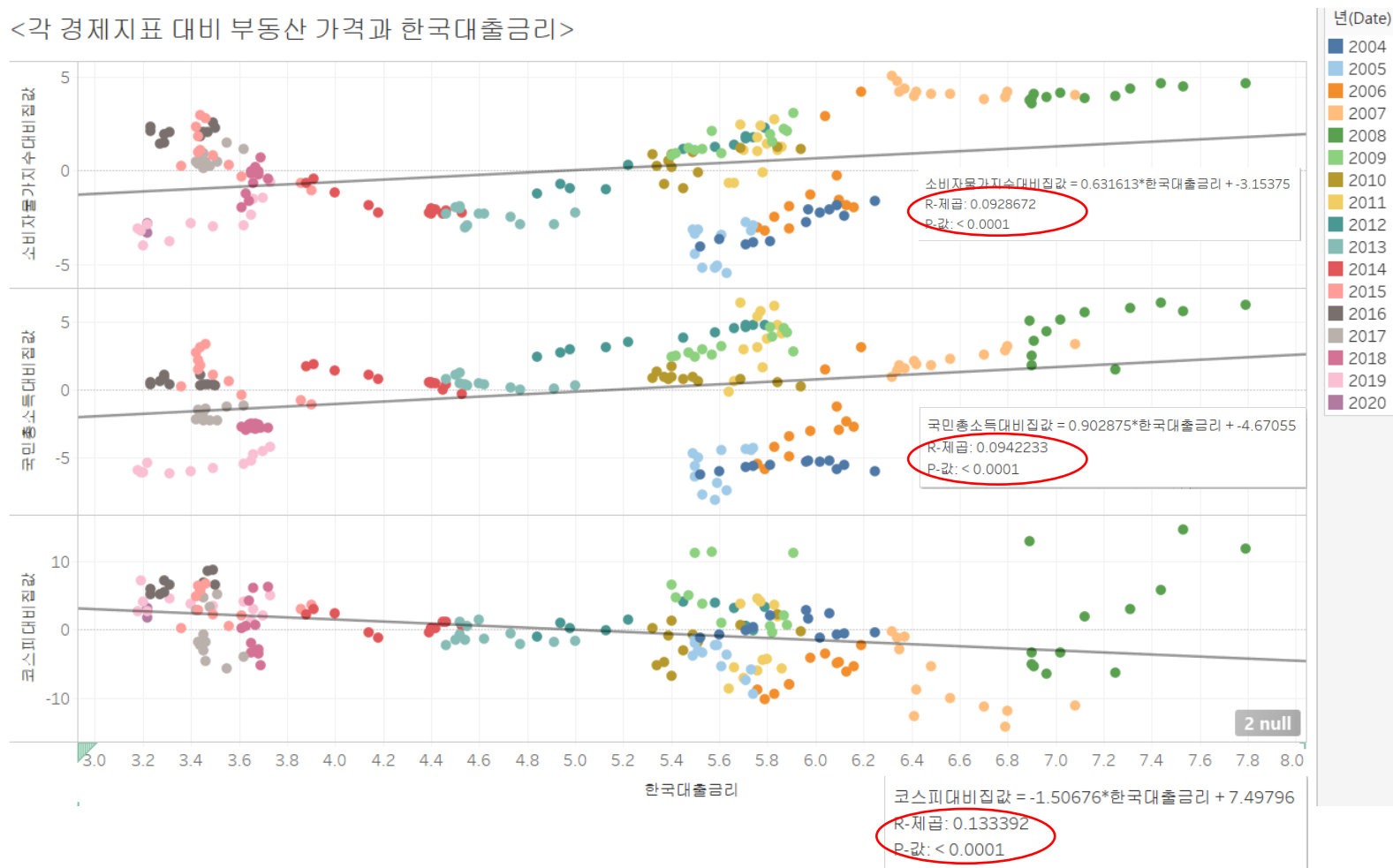
03 현행 분석

탐색적 분석을 통한 한국대출금리와 부동산 가격

한국대출금리와 각 경제지표 대비 전국 부동산 가격

<그림 7> 한국대출금리와 각 경제지표 대비 전국 부동산 판매 가격 데이터 scatter plot

<각 경제지표 대비 부동산 가격과 한국대출금리>



- 앞장의 $y - \hat{y}$ 을 한 경제지표대비 전국 부동산 가격의 Residual데이터와 한국대출금리 데이터를 scatter plot으로 나타냈다.
- 여전히 경제지표대비 전국 부동산 가격의 Residual 데이터와 한국대출금리 데이터를 선형식으로 표현하기엔 모델의 설명력이 부족하다.

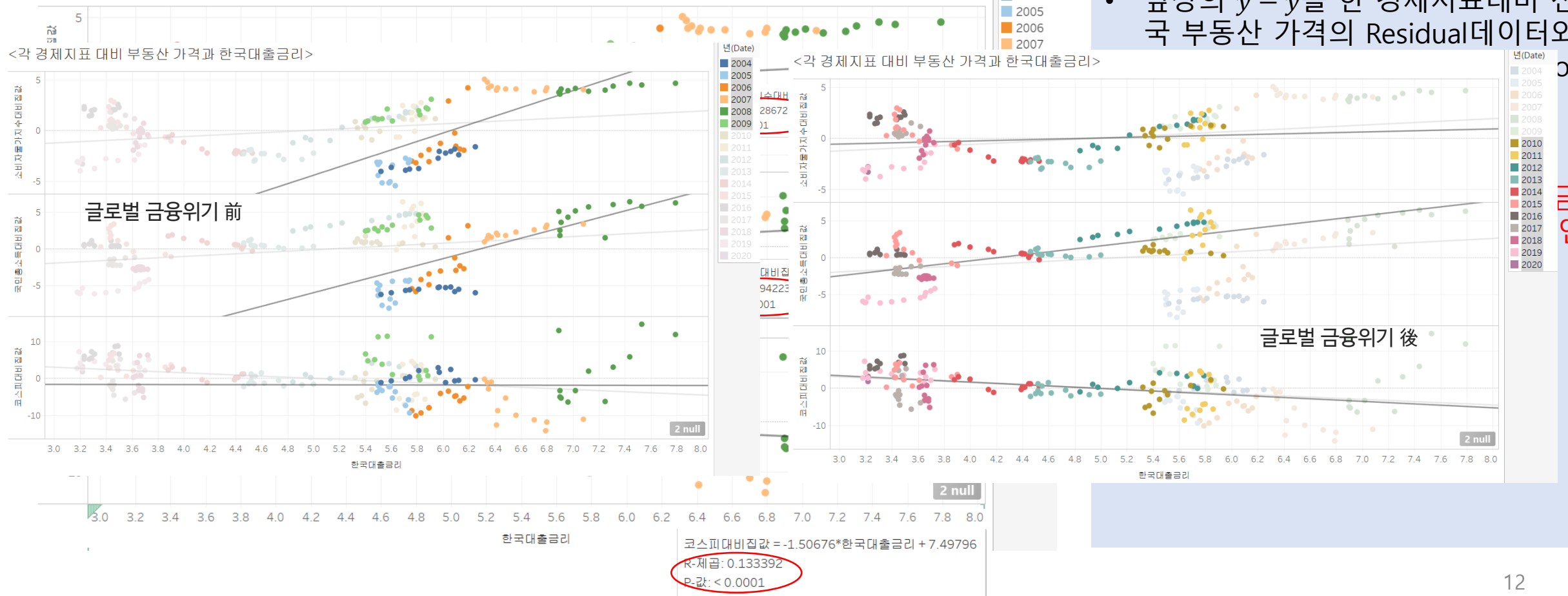
03 현행 분석

탐색적 분석을 통한 한국대출금리와 부동산 가격

한국대출금리와 각 경제지표 대비 전국 부동산 가격

<그림 7> 한국대출금리와 각 경제지표 대비 전국 부동산 판매 가격 데이터 scatter plot

<각 경제지표 대비 부동산 가격과 한국대출금리>



- 앞장의 $y - \hat{y}$ 을 한 경제지표대비 전국 부동산 가격의 Residual데이터와

가
금
엔

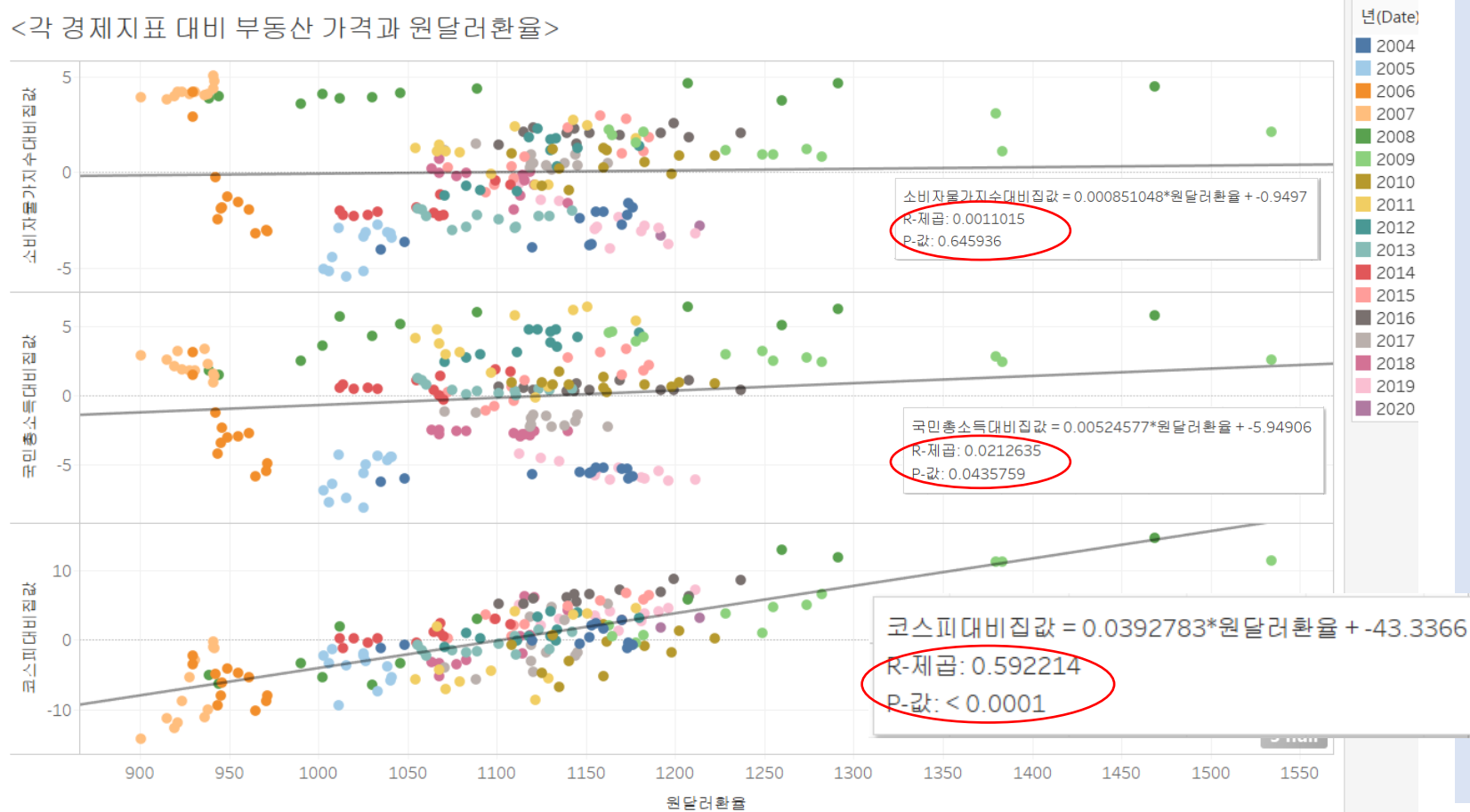
03 현행 분석

탐색적 분석을 통한 원달러환율과 부동산 가격

원달러환율과 각 경제지표 대비 전국 부동산 가격

<그림 8> 원달러환율과 각 경제지표 대비 전국 부동산 판매 가격 데이터 scatter plot

<각 경제지표 대비 부동산 가격과 원달러환율>



- 앞장의 $y - \hat{y}$ 을 한 경제지표대비 전국 부동산 가격의 Residual데이터와 원달러환율 데이터를 scatter plot으로 나타냈다.
- 코스피대비 전국 부동산 가격 데이터와 원달러환율은 양(+)의 상관관계를 가진다.
- 코스피대비 부동산 가격 데이터와 원달러환율 Linear Regression 결과의 R-square값이 0.592214로 모델이 데이터들을 잘 설명하고 있다.

03 현행 분석

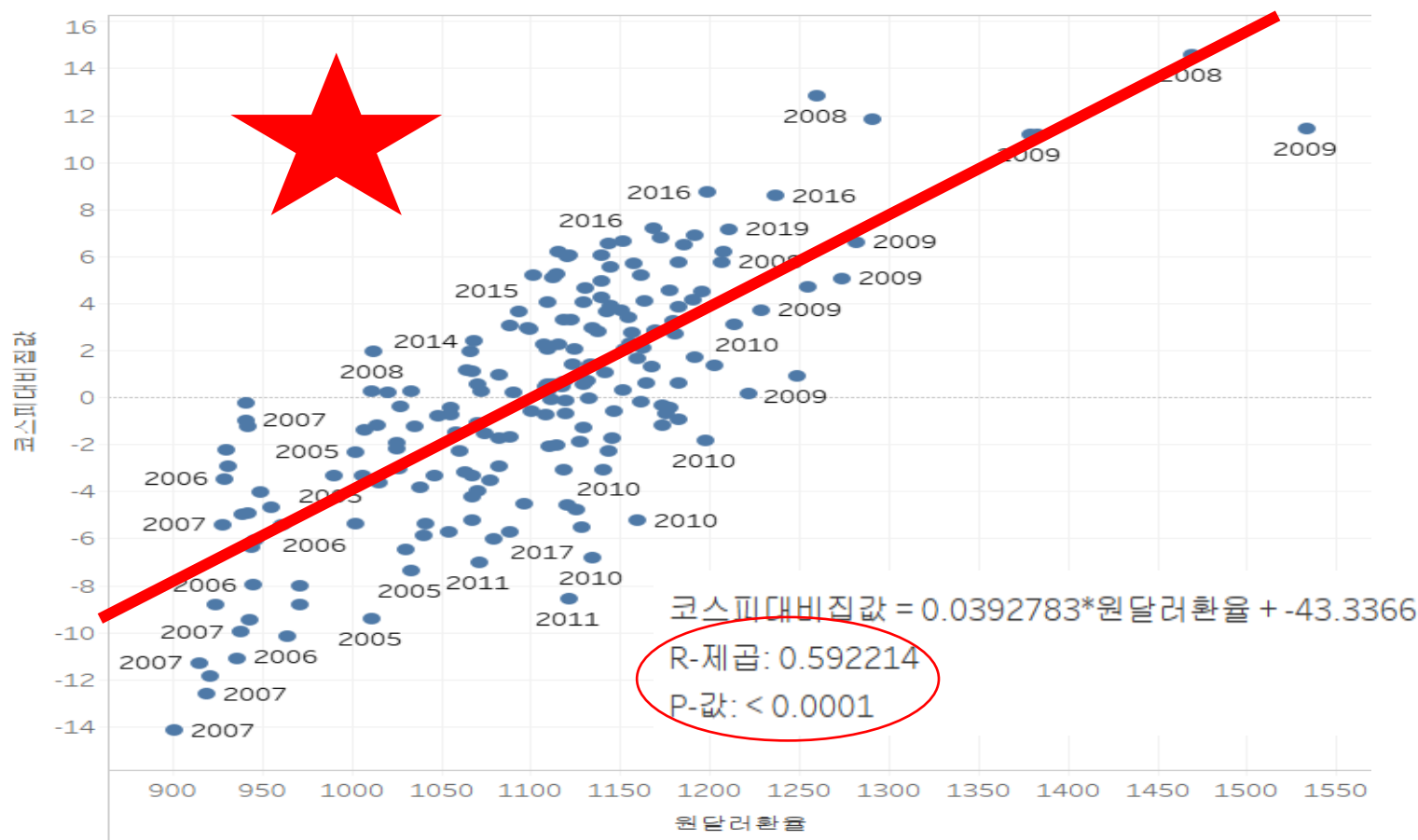
탐색적 분석을 통한 원달러환율과 코스피 대비 부동산 가격

가설 : 원 달러 환율이 오르면 부동산 가격이 하락한다.

원달러환율과 코스피 대비 전국 부동산 가격

〈그림 9〉 원달러환율과 코스피 대비 전국 부동산 판매 가격 데이터 scatter plot

<코스피 대비 집값과 원달러환율>



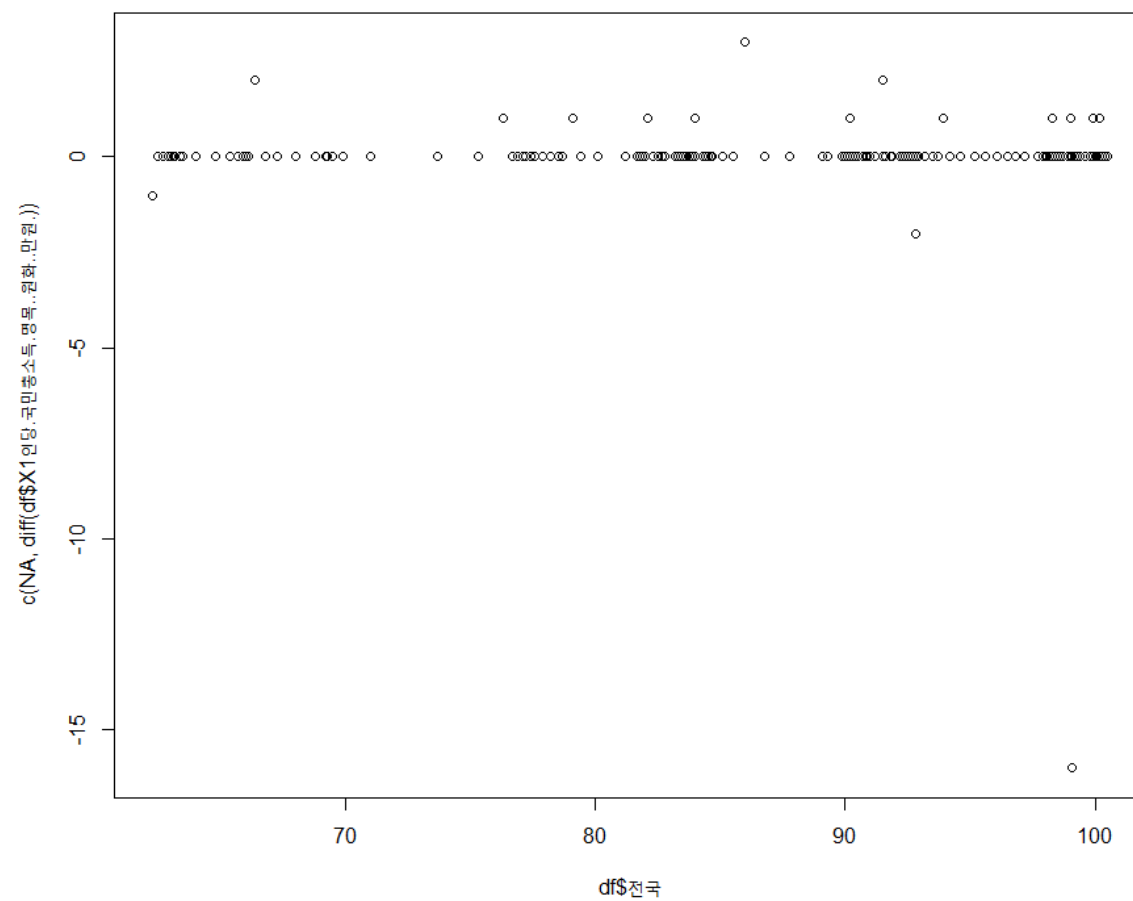
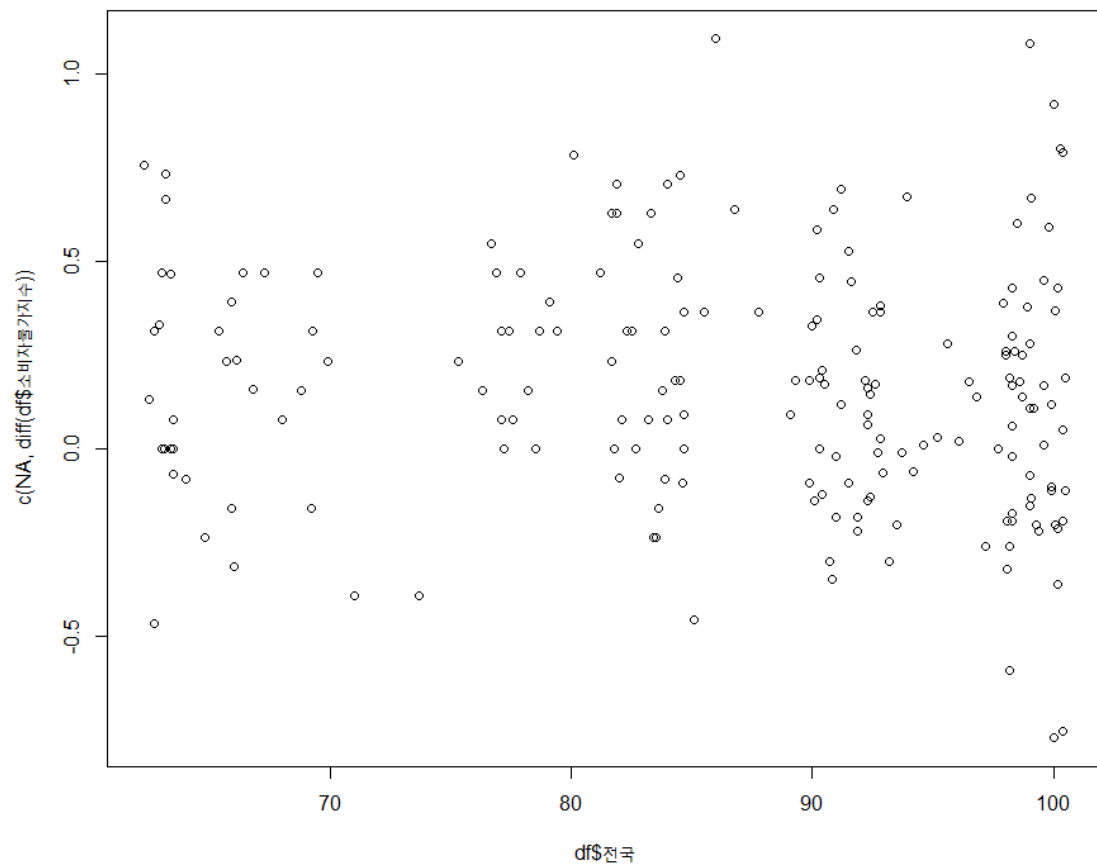
- 코스피대비 전국 부동산 가격 데이터와 원달러환율은 양(+)의 상관관계를 가진다.
- 코스피대비 부동산 가격 데이터와 원달러환율 Linear Regression 결과의 R-square값이 0.592214로 모델이 데이터들을 잘 설명하고 있다.
- 원달러환율이 오르면 부동산 가격이 하락한다는 가설을 기각한다.

03 현행 분석

탐색적 분석 diff

diff

〈그림 n〉 diff(경제지표) 전국 부동산 판매 가격 데이터 scatter plot

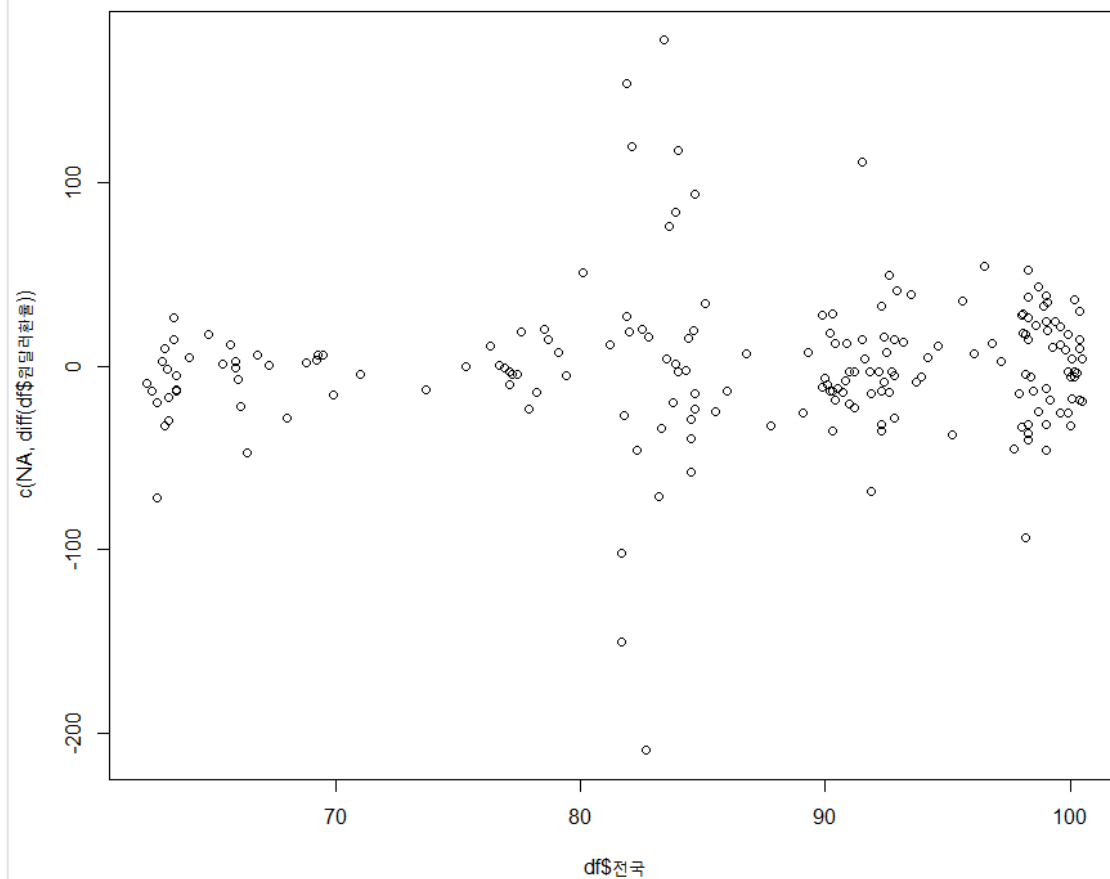
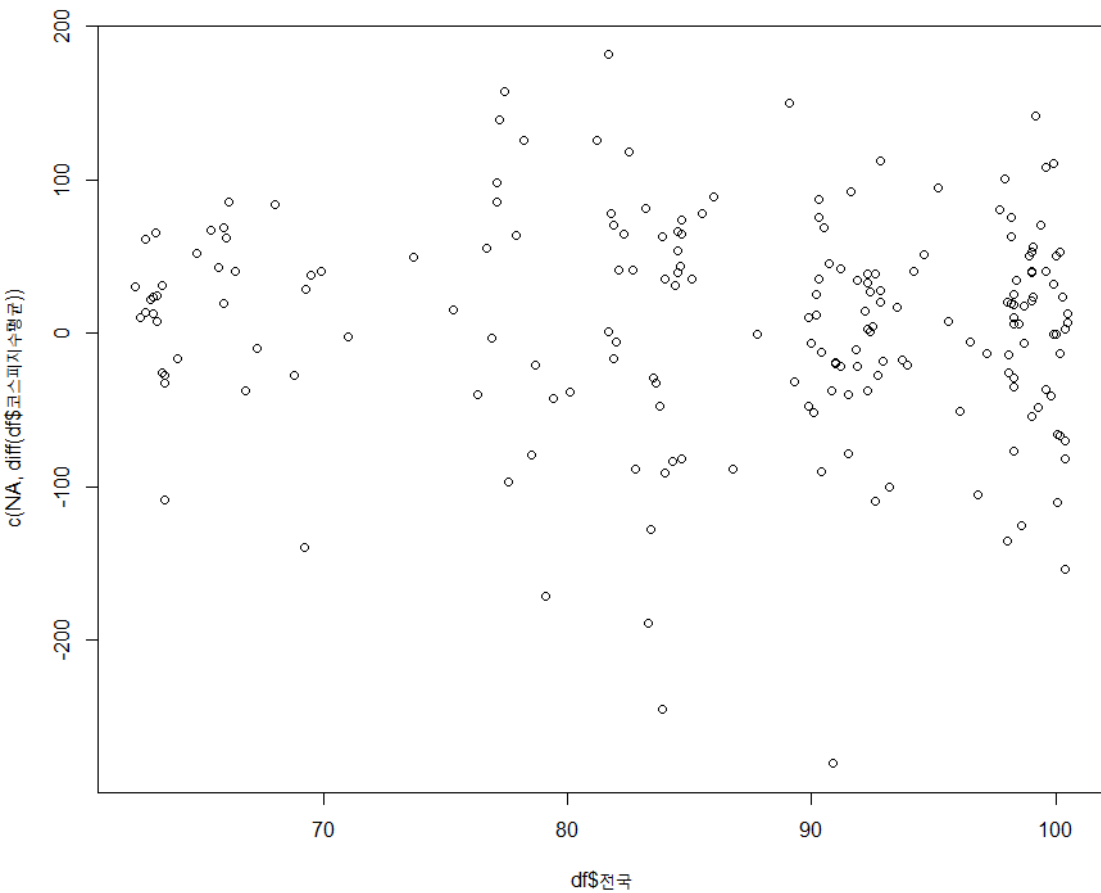


03 현행 분석

탐색적 분석 diff

diff

〈그림 n〉 diff(경제지표) 전국 부동산 판매 가격 데이터 scatter plot

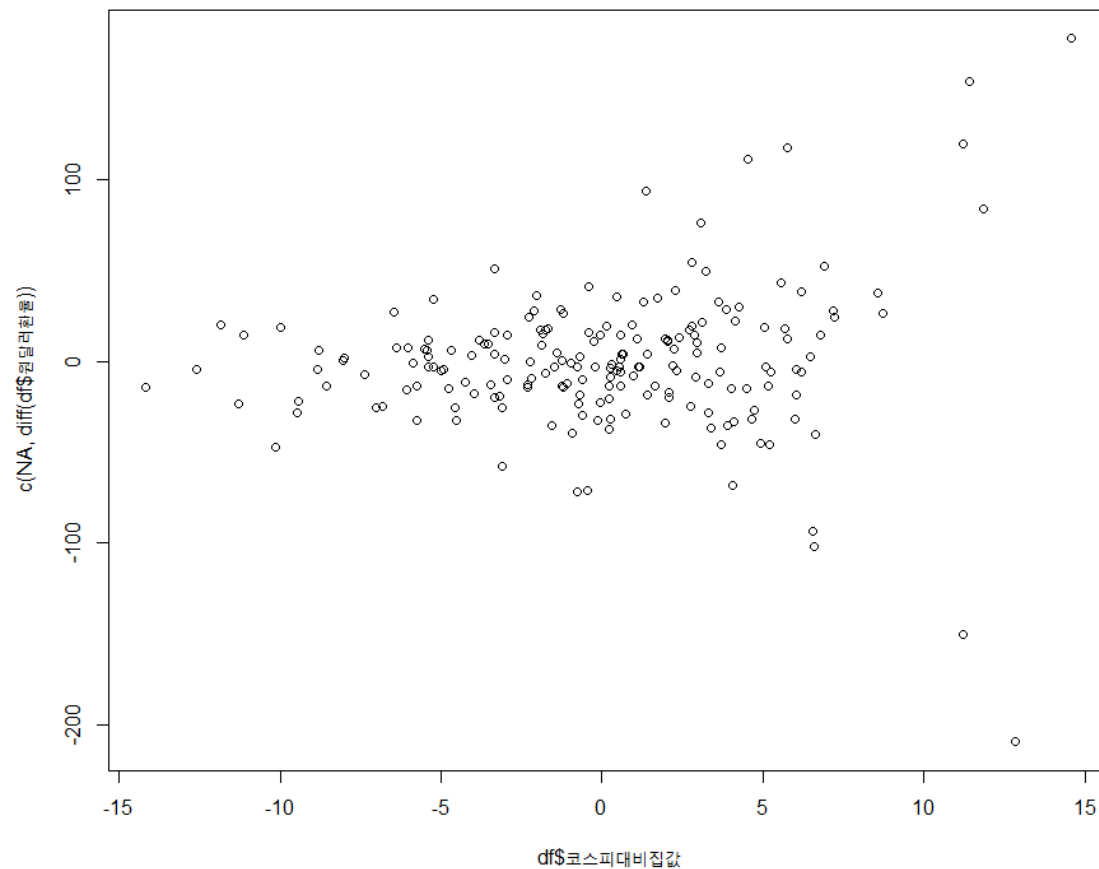
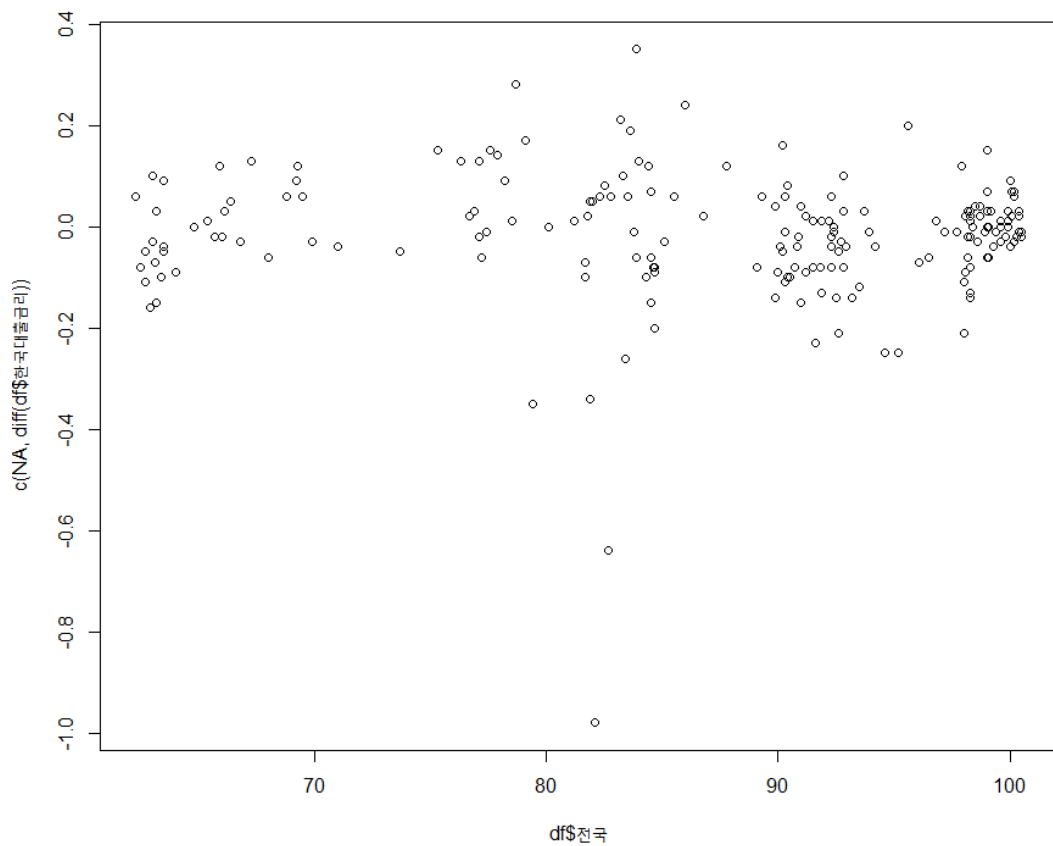


03 현행 분석

탐색적 분석 diff

diff

〈그림 n〉 diff(경제지표) 전국 부동산 판매 가격 데이터 scatter plot

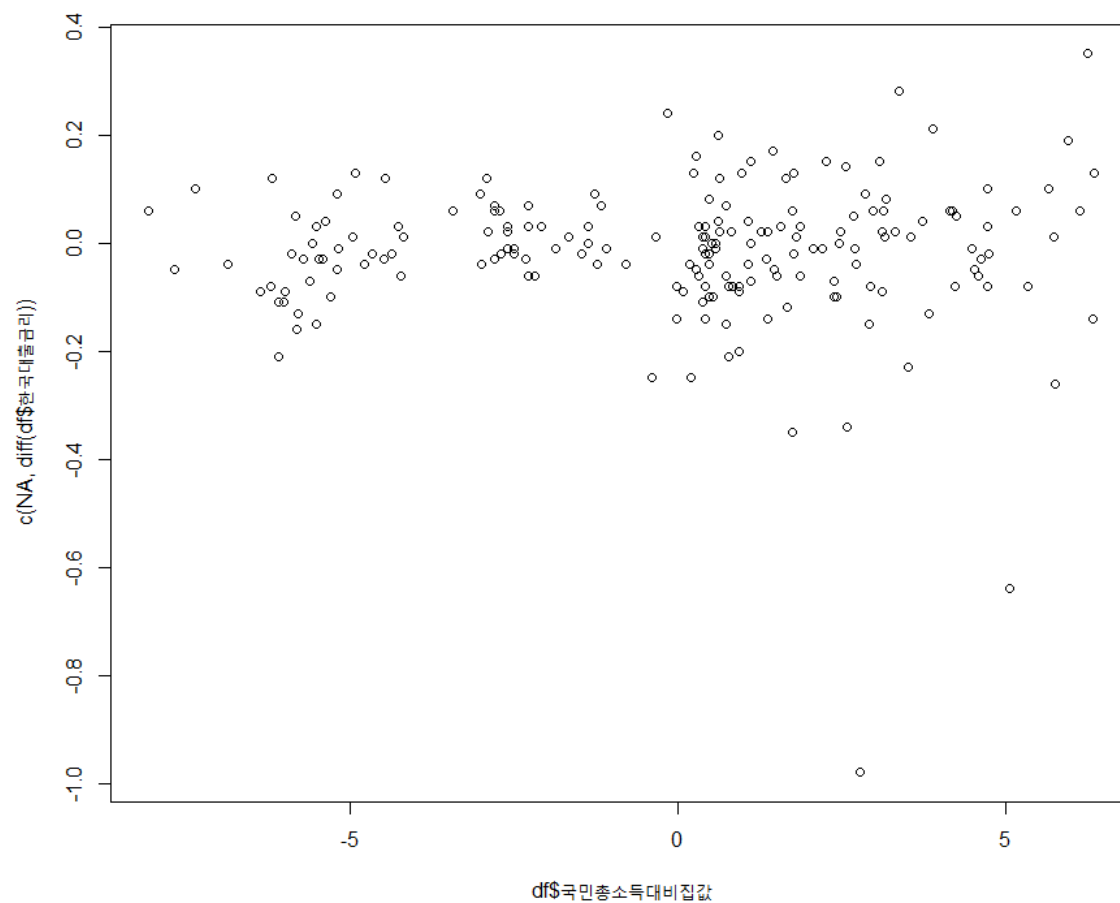


03 현행 분석

탐색적 분석 diff

diff

〈그림 n〉 diff(경제지표) 전국 부동산 판매 가격 데이터 scatter plot

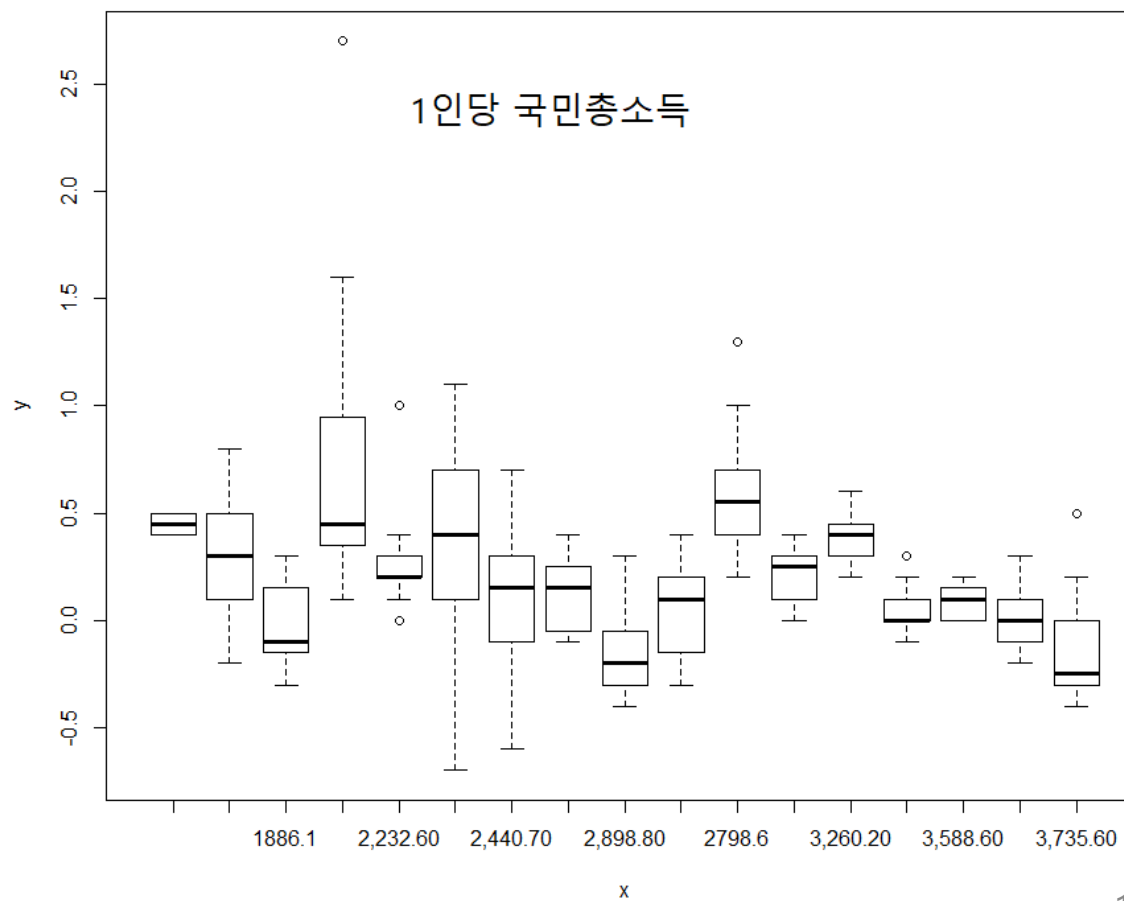
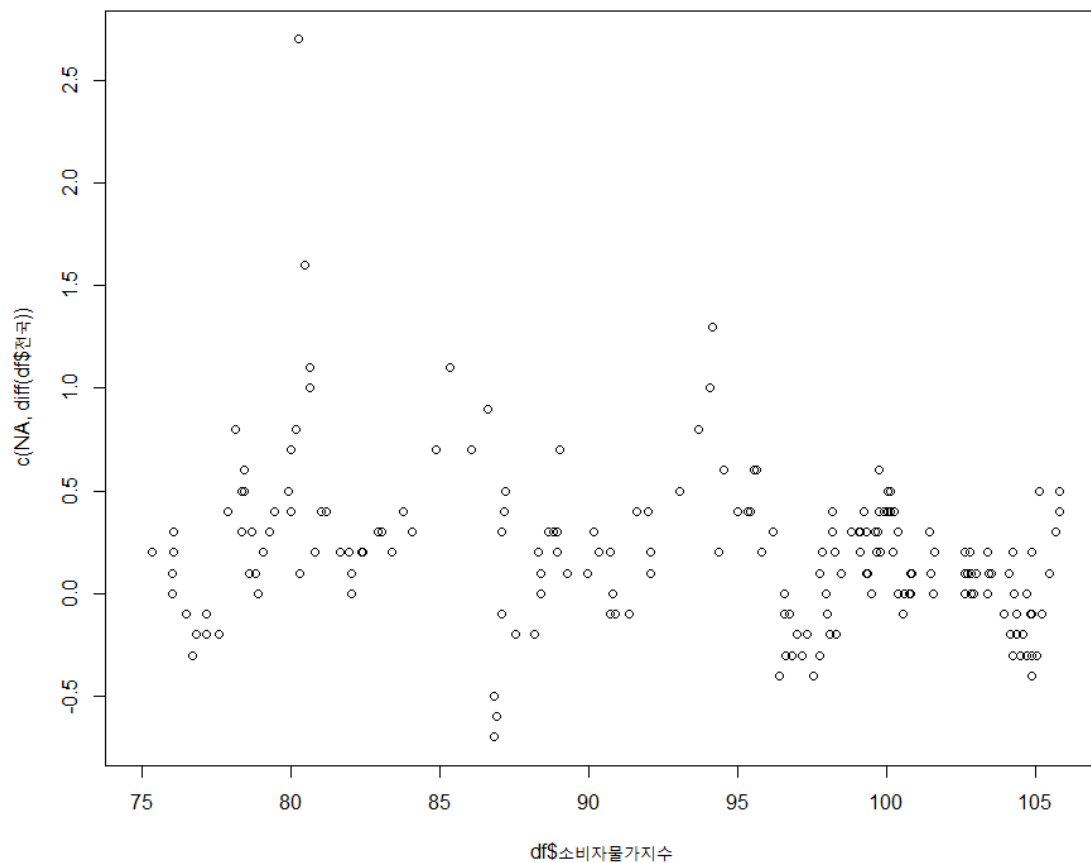


03 현행 분석

탐색적 분석 diff

diff

〈그림 n〉 경제지표 diff(전국 부동산 판매 가격) 데이터 scatter plot

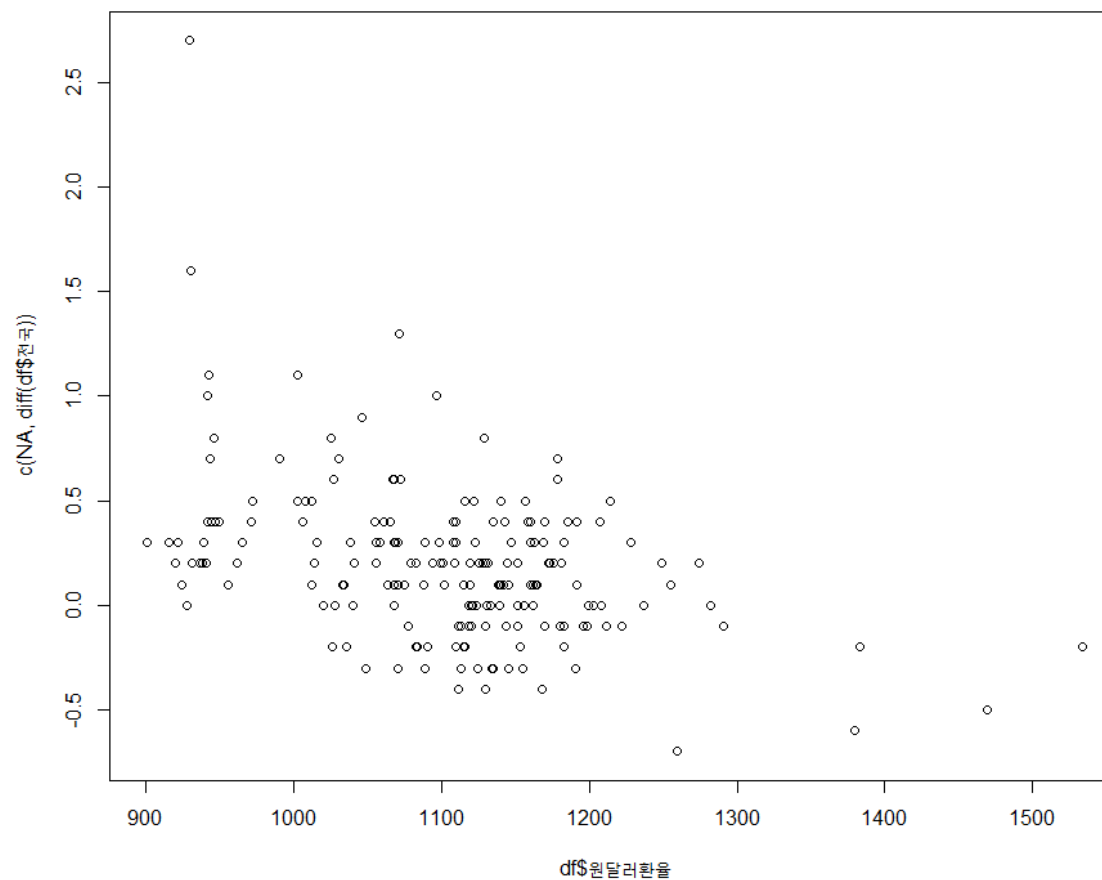
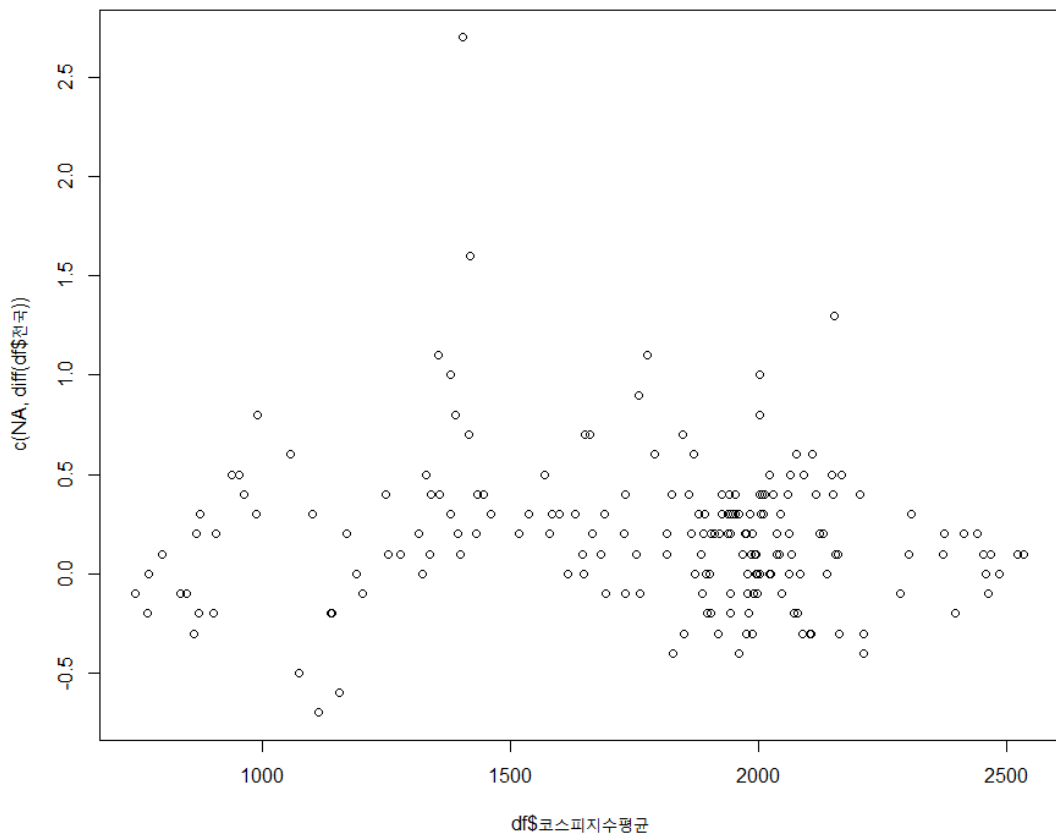


03 현행 분석

탐색적 분석 diff

diff

〈그림 n〉 경제지표 diff(전국 부동산 판매 가격) 데이터 scatter plot

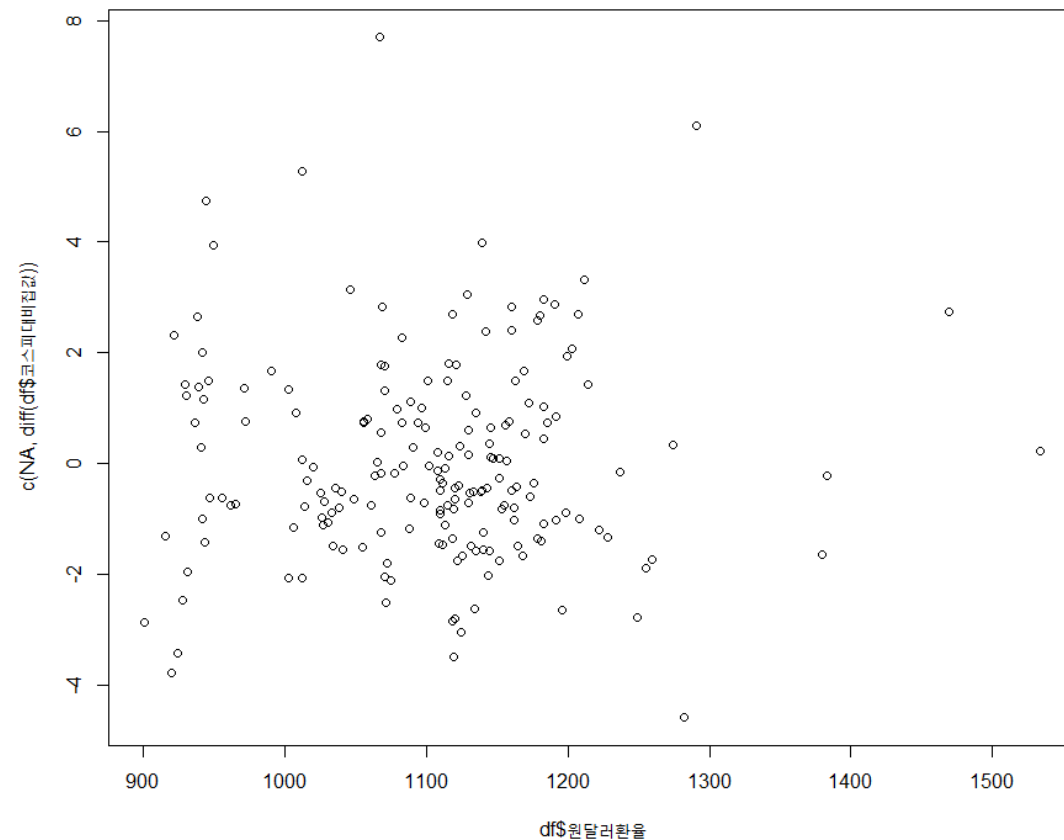
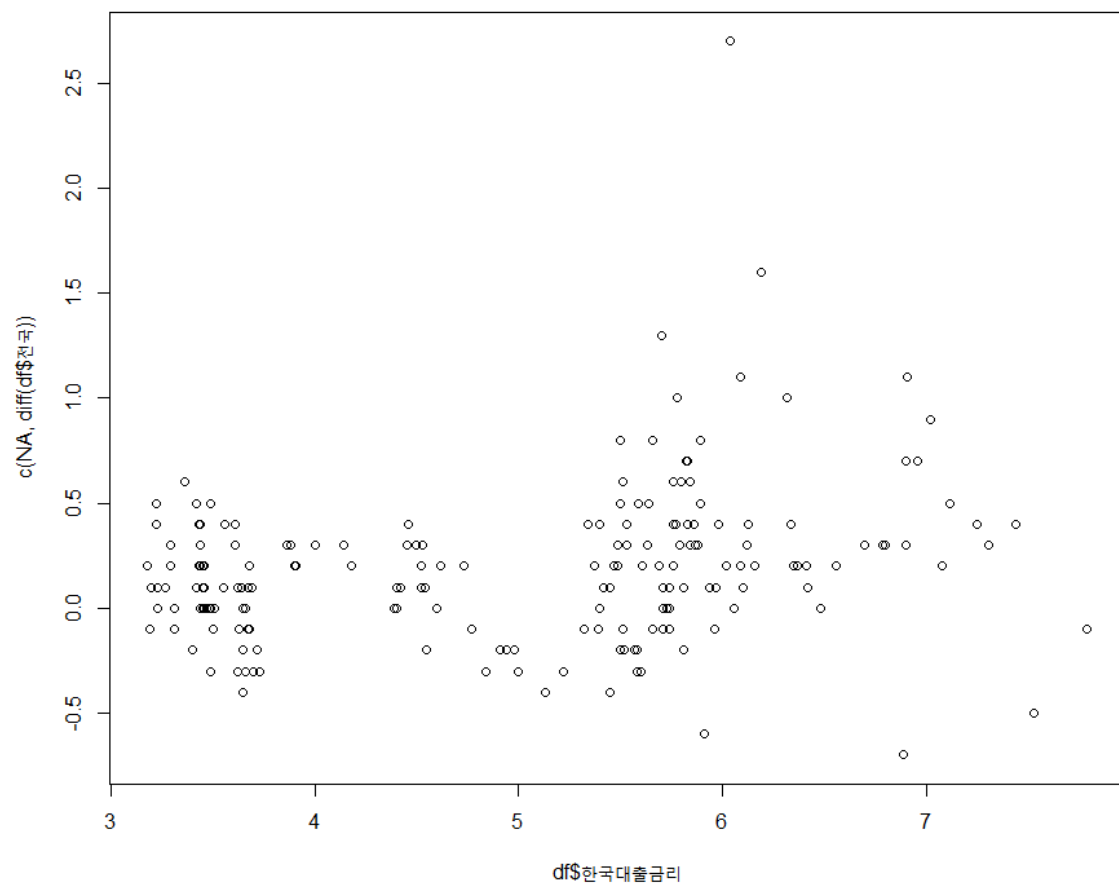


03 현행 분석

탐색적 분석 diff

diff

〈그림 n〉 경제지표 diff(전국 부동산 판매 가격) 데이터 scatter plot

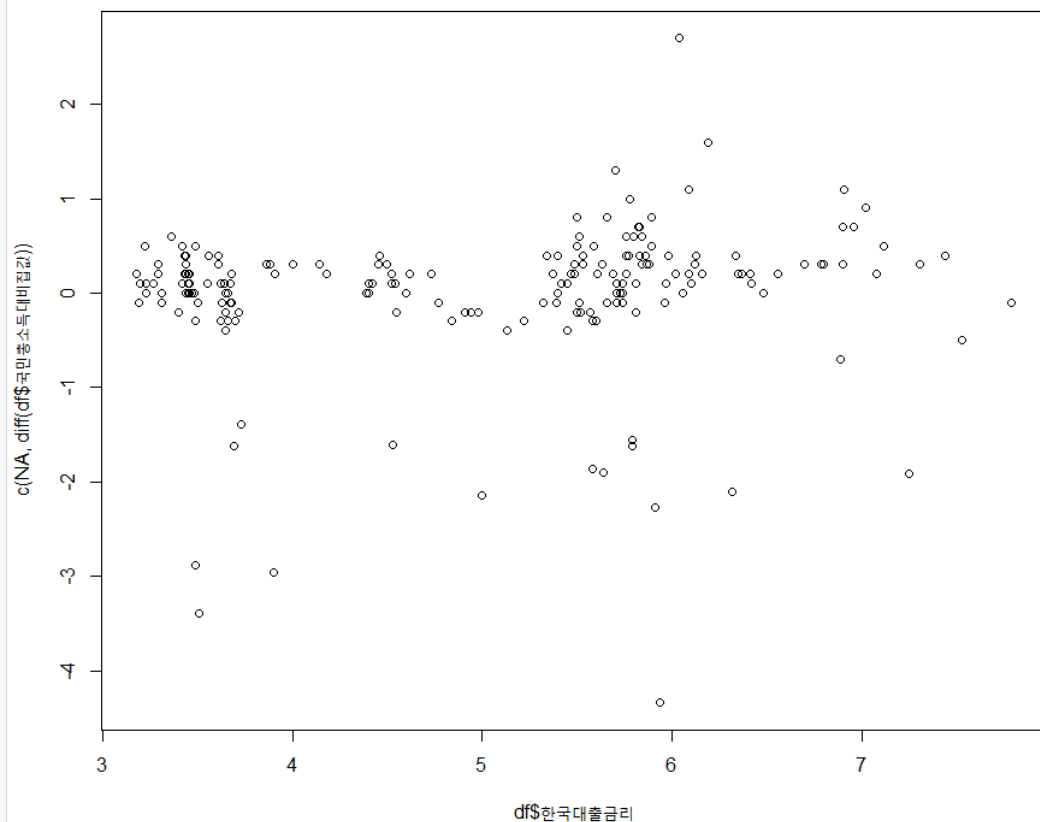


03 현행 분석

탐색적 분석 diff

diff

<그림 n> 경제지표 diff(전국 부동산 판매 가격) 데이터 scatter plot

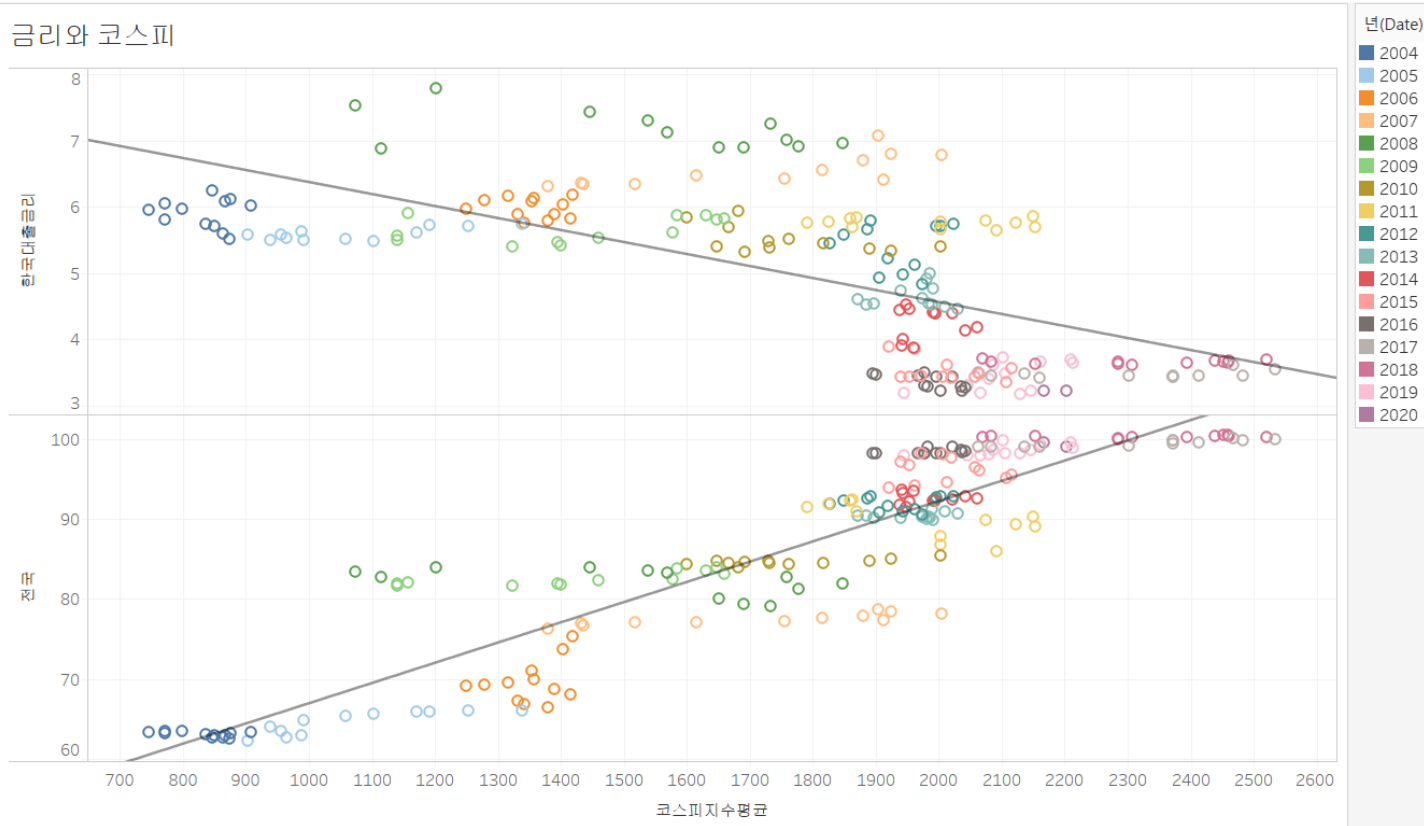


	diff 전국	diff코스피대비집값	diff국민대출금리	diff국민대출금리	diff소비자물가지수	diff코스피지수평균	diff원달러환율	diff한국대출금리
1	0.2	-0.35451901	0.200000	0	0.332	21.92	2.5	-0.16
2	0.3	0.09433214	0.300000	0	0.732	8.13	-29.6	0.03
3	0.2	-0.59836594	0.200000	0	0.000	31.56	26.7	-0.10
4	0.1	2.84097329	0.100000	0	-0.066	-108.35	-13.2	-0.05
5	0.0	0.68758333	0.000000	0	0.000	-27.18	-4.6	0.09
6	-0.1	0.54609560	-0.100000	0	0.466	-25.54	14.5	-0.10
7	-0.2	-0.81320898	-0.200000	0	0.665	24.24	-17.0	-0.15
8	-0.1	-1.75976241	-0.100000	0	0.000	65.61	-1.2	-0.07
9	-0.1	-0.43367271	-0.100000	0	0.000	13.19	-32.2	-0.03
10	-0.3	-0.65391063	-0.300000	0	-0.465	13.99	-71.4	-0.11
11	-0.2	-0.45196210	-0.200000	0	0.133	9.96	-13.1	-0.08
12	-0.2	-0.96625825	-1.862258	-1	0.756	30.29	-9.5	0.06
13	0.4	-1.14643006	0.400000	0	0.315	61.13	-19.6	-0.05
14	0.3	-0.30081325	0.300000	0	0.470	23.75	9.5	0.10
15	0.5	1.33026067	0.500000	0	0.079	-32.82	-12.7	-0.04
16	0.5	0.91563628	0.500000	0	-0.079	-16.43	4.9	-0.09
17	0.3	0.58127165	0.300000	0	0.335	53.35	17.8	0.00
18	-0.1	-2.65427848	-0.100000	0	0.390	100.97	-15.0	0.12
189	0.1	-0.41960860	0.100000	0	0.260	20.54	-32.8	-0.11
190	0.2	-1.39474810	0.200000	0	-0.590	63.04	17.8	-0.02
191	0.5	0.03908137	0.500000	0	0.250	18.22	-24.8	0.04
192	0.4	-1.02753228	NA	-16	0.670	56.43	35.4	0.00
193	0.5	1.41880157	NA	0	0.010	-36.32	21.9	0.00
194	NA	NA	NA	0	NA	NA	NA	NA
195	NA	NA	NA	0	NA	NA	NA	NA

03 현행 분석

탐색적 분석 time lag

Time Lag 주기 전에,,



- 대출금리 상승
→ 주식 투자자 증가

- 코스피지수 상승 → 비유동자산(부동산) 투자 심리 상승 → 부동산 가격 상승

On 연구 과정

데이터
수집

주요 Data
부동산 가격,
원 달러 환율,
한국 대출금리,
국민총소득,
코스피 지수 평균

예측
모델 수립

사용 툴
R studio

데이터
전처리

사용 툴
Excel

모델 검증

Training Data : Test Data
7:3

5월 중순

현행분석
EDA를
통한 변수
정의

사용 툴
Tableau,
Minitab14,
R studio

의사결정

0n 참고 문헌

김기환, "집 값 올라서 이혼 늘었을까요", <중앙일보> 2017.05.15, 7면

김남현, 이근영. 2016. *금리와 주택가격*. p.45

김남현, 장한익. 2018. *금리가 주택가격에 미치는 영향과 요인*. 주택금융연구원, p.7

김윤영. 2012. *우리나라 주택시장의 매매. 전세 가격변동 거시결정요인의 동태분석*. P.127

장영길 (Chang Young-gil). 2017. *저금리가 주택가격을 상승시키는가?*. 부동산학연구, 23(1): 39-48

한국 금융 연구원. 2008. *글로벌 금융위기의 영향 및 전망*

감사합니다
