

R D D / D I D / S A R I M A

담배값 인상과 성인 흡연율



5조

2020103907	강민영
2021103961	곽동엽
2020103910	권수영
2019100873	박건우
2019100890	이승재
2019100895	이윤서

들어가며



ews

11TH
ANNIVERSARY

여행·레저

건설·부동산

공기업·지자체

문화·라이프

포토·리뷰

역대급 세수 펑크에 '담뱃값 인상설', 한 갑 8000원 될까?

"총선 후 담뱃값 오른다?"...한 갑 4500원→8000원·1만원
'인상설 솔솔'

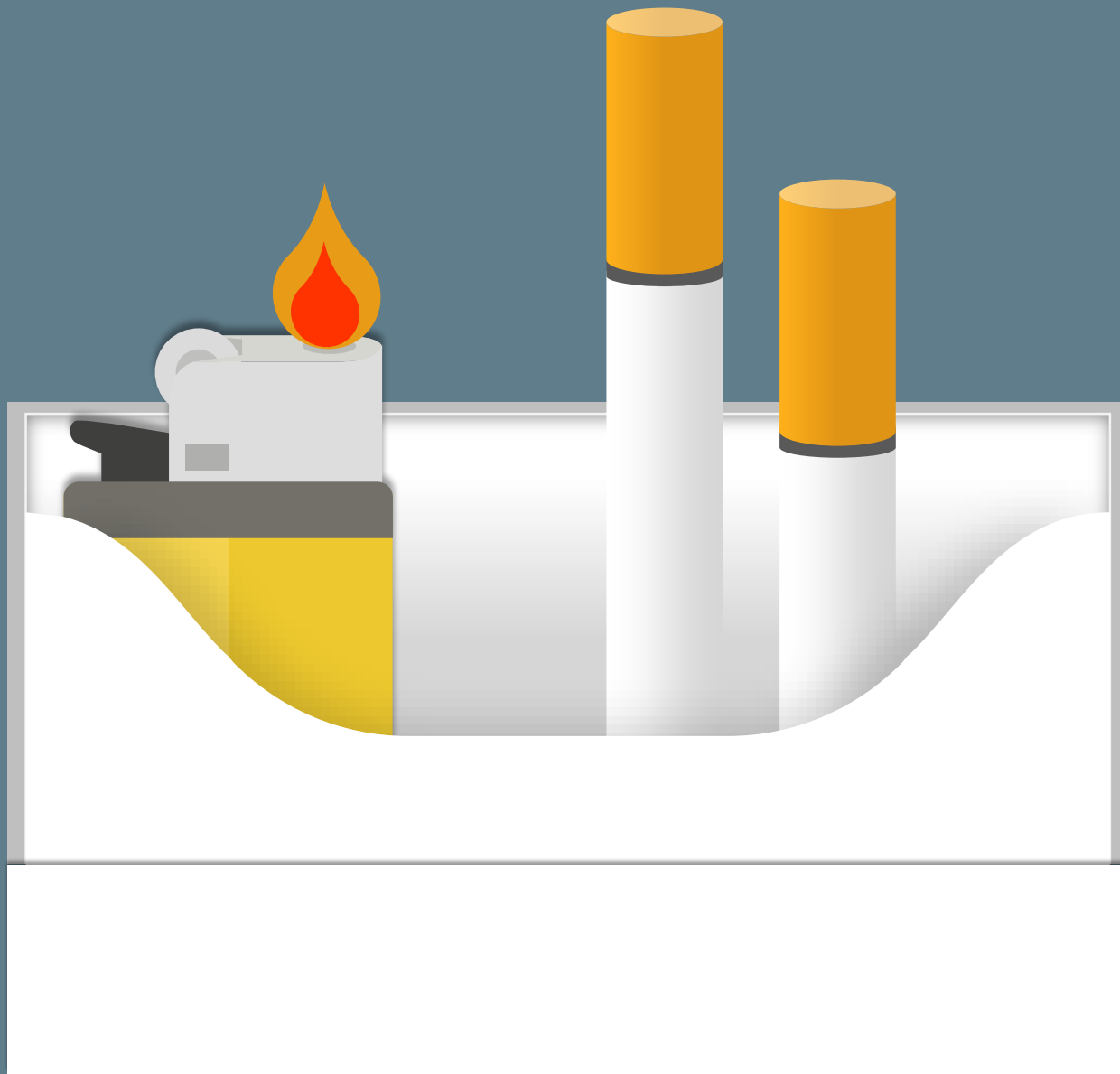
조선경제 > 산업·재계

담뱃값 8000원 시대 온다? "내년 총선 이후 인상 예상"

담뱃값 인상과 흡연율은 인과성이 있을까?
담뱃값을 인상하면 실제적인 흡연율 감소 효과가 있을까?

인과성 분석

RDD



인과성 분석(RDD) 개요 및 데이터 수집

분석 목적 : 2015년 담뱃값 인상이 한국 성인 흡연에 미친 단기적 영향 분석

처치:

2015년 담뱃값 인상 (2500원 → 4500원)

결과 변수:

월간 담배 판매량

할당 변수:

시간 (월 단위)

임계값:

2015년 1월

수집 기간:

2012년 1월 ~ 2016년 12월

월간 담배 판매량

출처:

기획재정부, '2022년 상반기 담배 시장 동향'

한국담배협회, '2012년 1월 ~ 2015년 7월 월별 담
배 판매량'

흡연율 > 월간 판매량 대체 타당성 검증

흡연율 연 단위 조사

→ 월 단위 데이터 존재하는 담배 판매량 데이터 대체

상관계수 :

연도	현재 흡연율	담배판매량 합계(단위: 억갑)	매일 흡연율
2012	25.80%	43.6	23.60%
2013	24.10%	43.2	21.40%
2014	24.20%	43.5	21.70%
2015	22.60%	33.2	18.70%
2016	23.90%	36.7	20.20%

선행연구 :

‘담뱃세 인상정책의 흡연억제 효과
- 시계열자료를 통한 실증분석’ 에서
담배판매량으로 흡연이 감소했는지 확인

현재 흡연율 - 담배 판매량 합계 간의 상관계수: 0.790603082

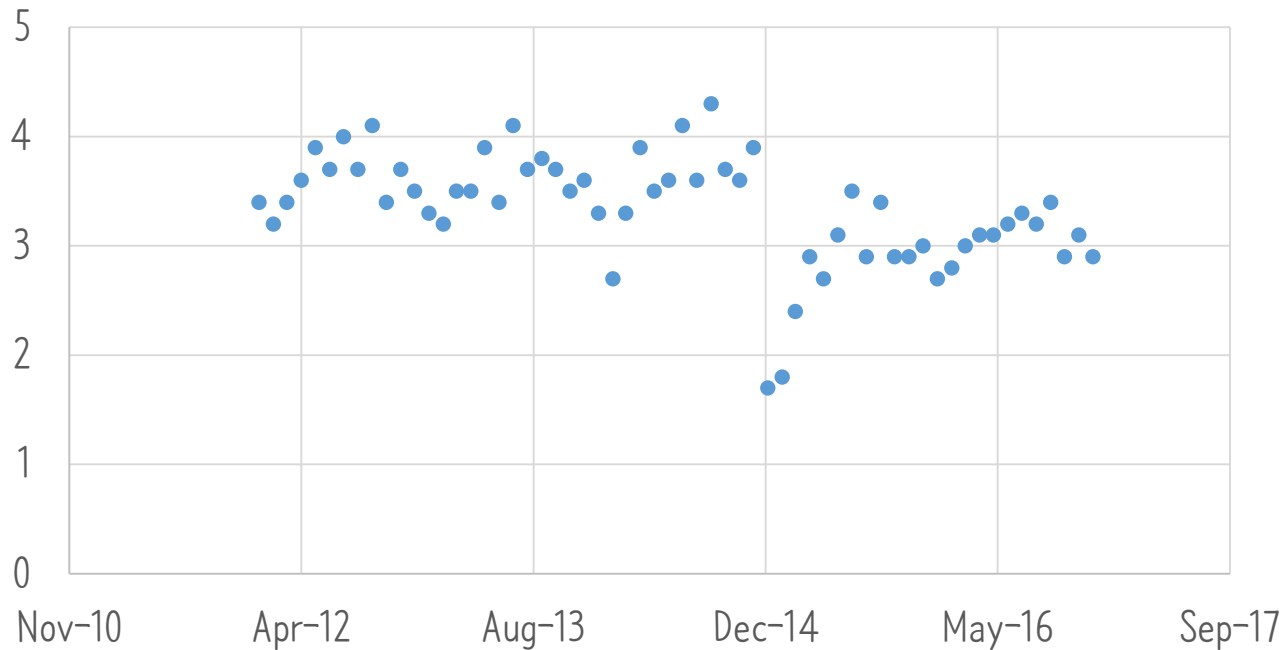
매일 흡연율 - 담배 판매량 합계 간의 상관계수: 0.894416031

RDD 가정 성립 확인

1) 할당 변수와 결과 변수의 연속성 :

임계값 기준 자연스러운 연속적인 분포

월별 담배 판매량



2) 조작 가능성의 부재

담뱃값 인상 정책 - 정부에 의해 결정 및 예고
정책 시점 조작 불가

3) 임계값 주변의 동질성

흡연자 특성: 담뱃값 인상 시점 전후
흡연자 집단의 특성 / 흡연 습관
큰 차이 X

외부 요인: 담뱃값 인상 시점 전후
다른 정책 변화 / 사회적 이슈 X

RDD 분석 - 대역폭 및 모형 선택

대역폭 선택:

- 1) 6개월
- 2) 1년
- 3) 1년 6개월

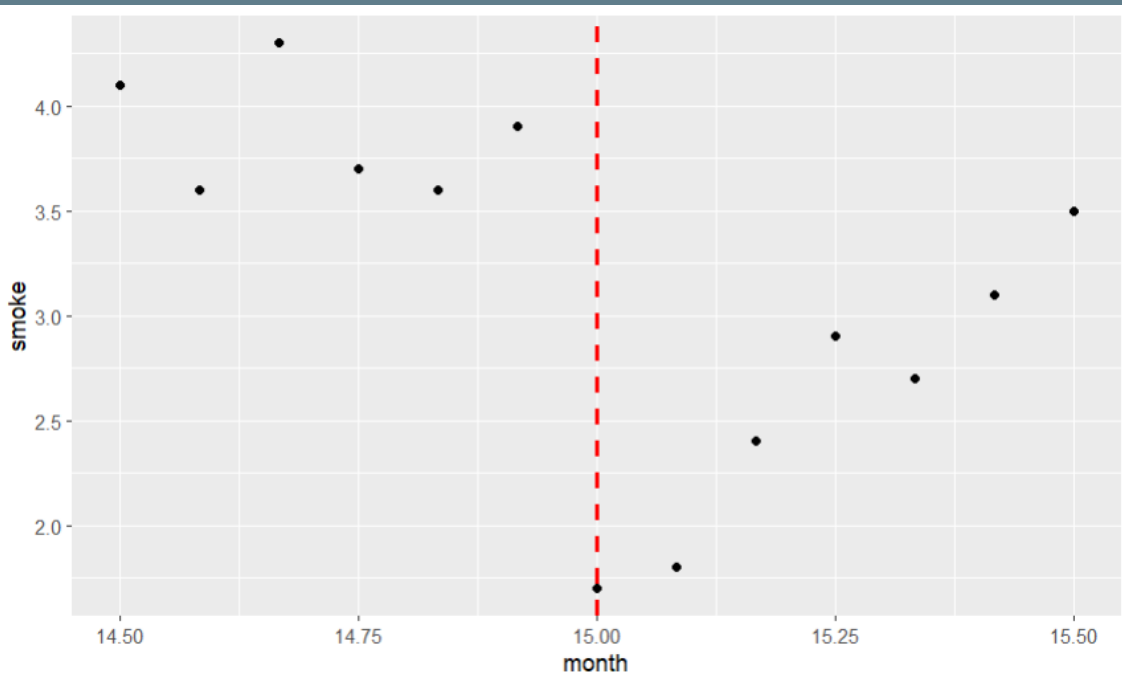
모형 선택

- 1) RDD 회귀모형
- 2) 비선형 RDD 회귀모형 (2차항)
- 3) 비선형 RDD 회귀모형 (3차항)

RDD 분석 - 대역폭 및 모형 선택

대역폭 6개월

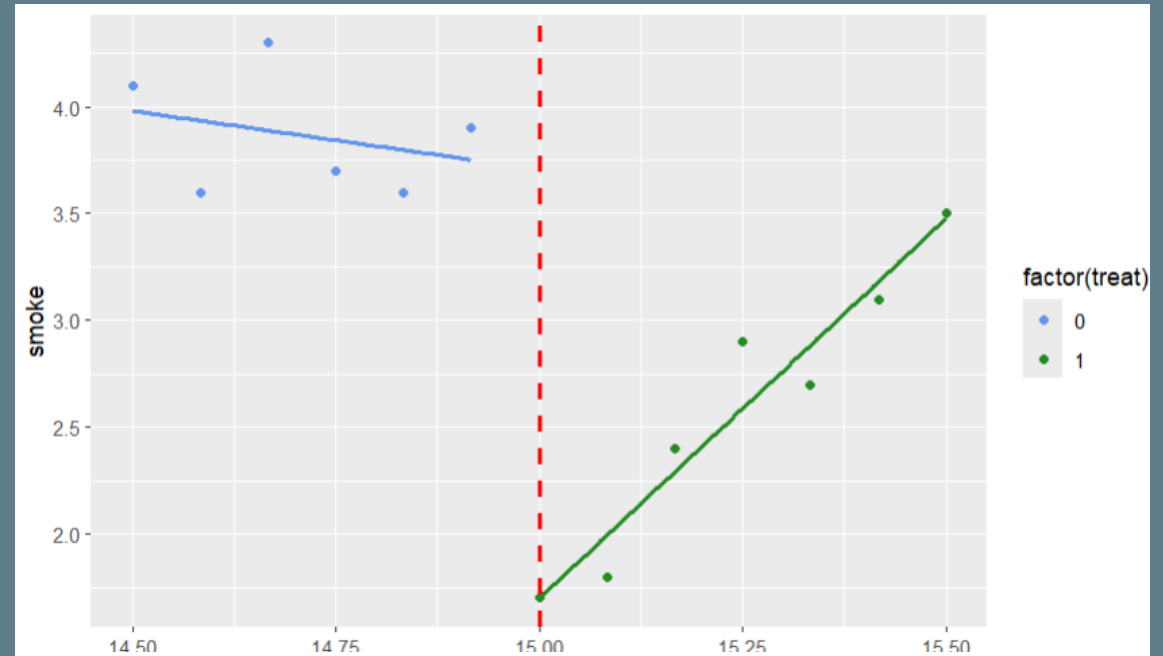
대역폭 6개월 기본 시각화



변수 목록

month: 년월 (개월 기준 소수점 표기)
increase: 담뭍값 인상 년월
month_d: 년월 - 담뭍값 인상 년월
treat: 담뭍값 인상 여부
smoke: 담배 판매량 (단위:억갑)

RDD 회귀모형 결과 시각화 및 모수 추정 결과



Coefficients:

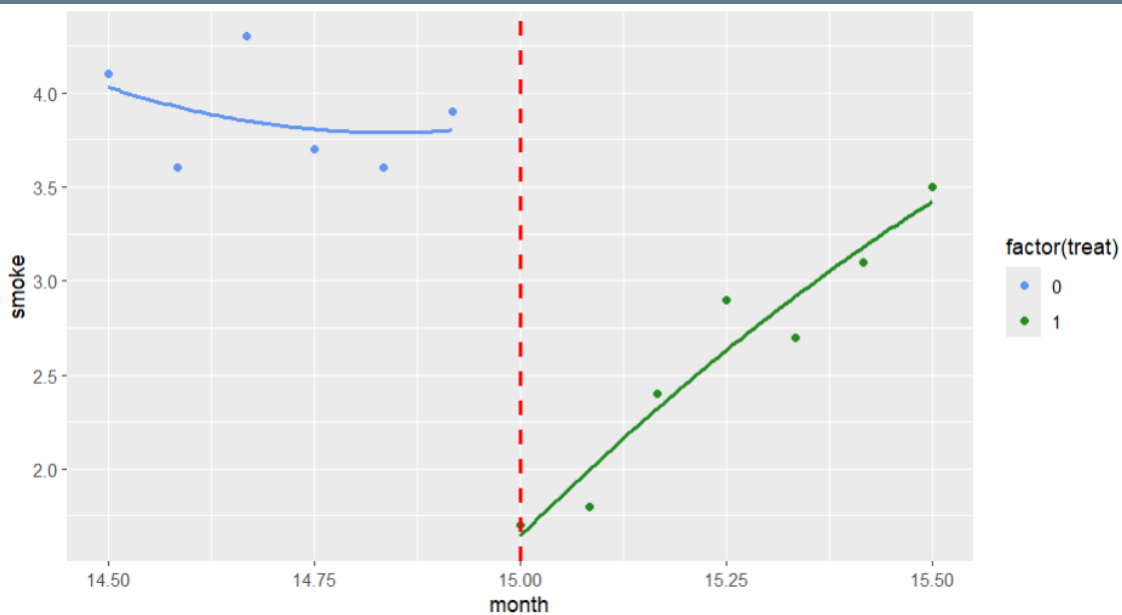
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	3.7067	0.2332	15.891	6.83e-08	***
treat	-2.0107	0.2891	-6.955	6.65e-05	***
month_d	-0.5486	0.7187	-0.763	0.46484	
treat:month_d	4.1057	0.9162	4.481	0.00153	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RDD 분석 - 대역폭 및 모형 선택

대역폭 6개월

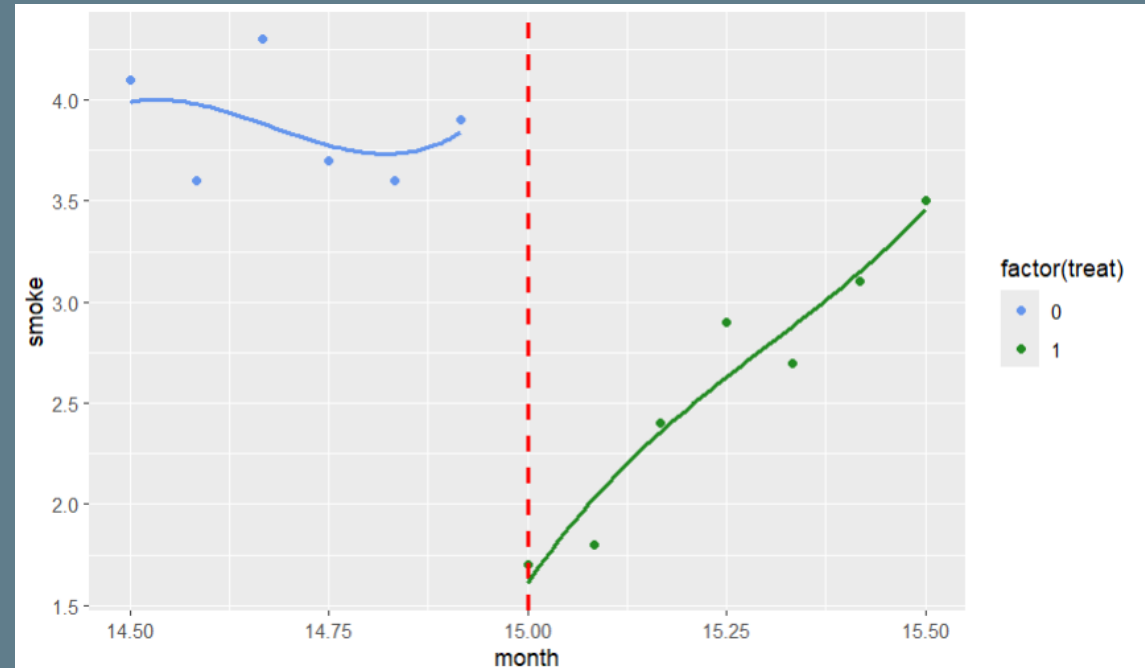
비선형 RDD 회귀모형(2차항) 결과 시각화 및 모수 추정 결과



Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	3.8400	0.5004	7.674	0.000119	***
treat	-2.1971	0.5568	-3.946	0.005559	**
month_d	0.6514	3.9284	0.166	0.872981	
I(month_d^2)	2.0571	6.5924	0.312	0.764093	
treat:month_d	3.6771	4.5457	0.809	0.445176	
treat:I(month_d^2)	-3.6000	7.9231	-0.454	0.663317	

비선형 RDD 회귀모형(3차항) 결과 시각화 및 모수 추정 결과



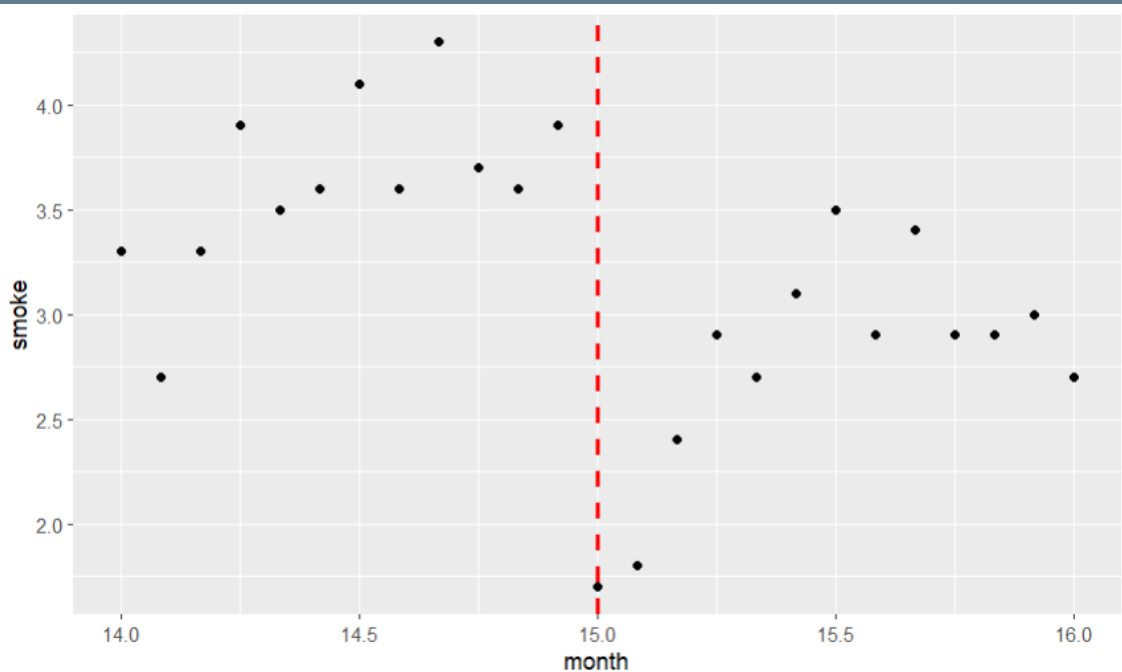
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	4.16667	1.17406	3.549	0.0164	*
treat	-2.55714	1.21527	-2.104	0.0893	.
month_d	5.58254	16.05275	0.348	0.7422	
I(month_d^2)	21.65714	61.64165	0.351	0.7397	
I(month_d^3)	22.40000	69.89952	0.320	0.7616	
treat:month_d	0.07937	17.11885	0.005	0.9965	
treat:I(month_d^2)	-30.40000	68.19362	-0.446	0.6744	
treat:I(month_d^3)	-12.80000	79.69771	-0.161	0.8787	

RDD 분석 - 대역폭 및 모형 선택

대역폭 12개월

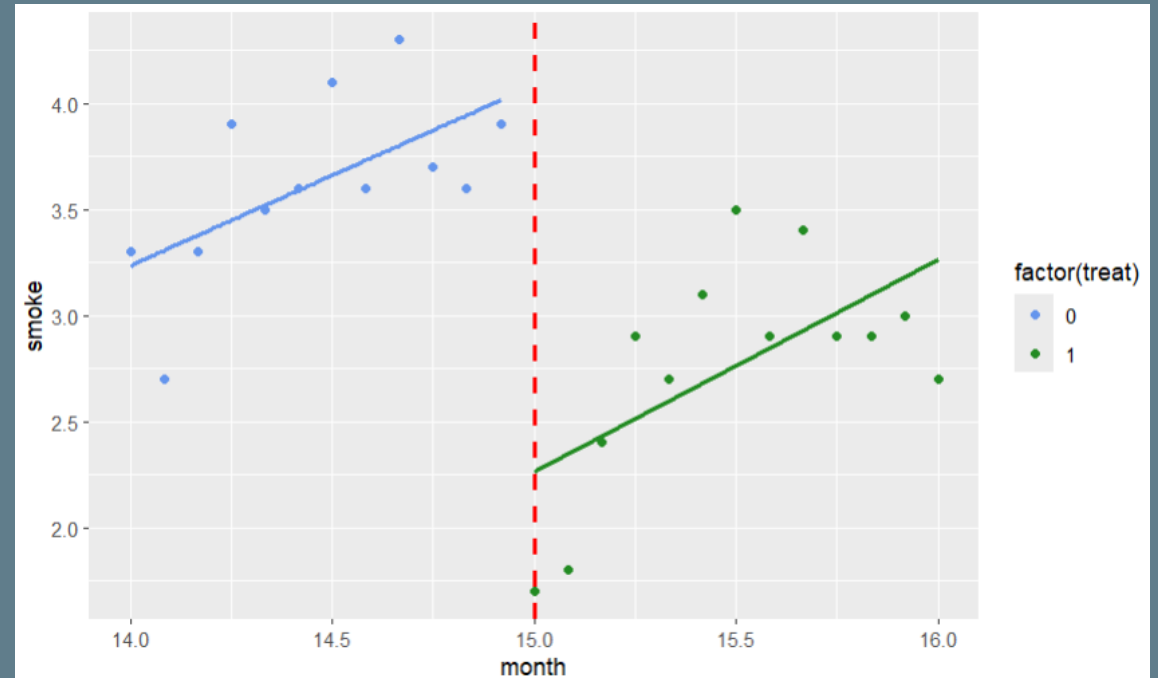
대역폭 6개월 기본 시각화



변수 목록

month: 년월 (개월 기준 소수점 표기)
increase: 담뭍값 인상 년월
month_d: 년월 - 담뭍값 인상 년월
treat: 담뭍값 인상 여부
smoke: 담배 판매량 (단위:억갑)

RDD 회귀모형 결과 시각화 및 모수 추정 결과



Coefficients:

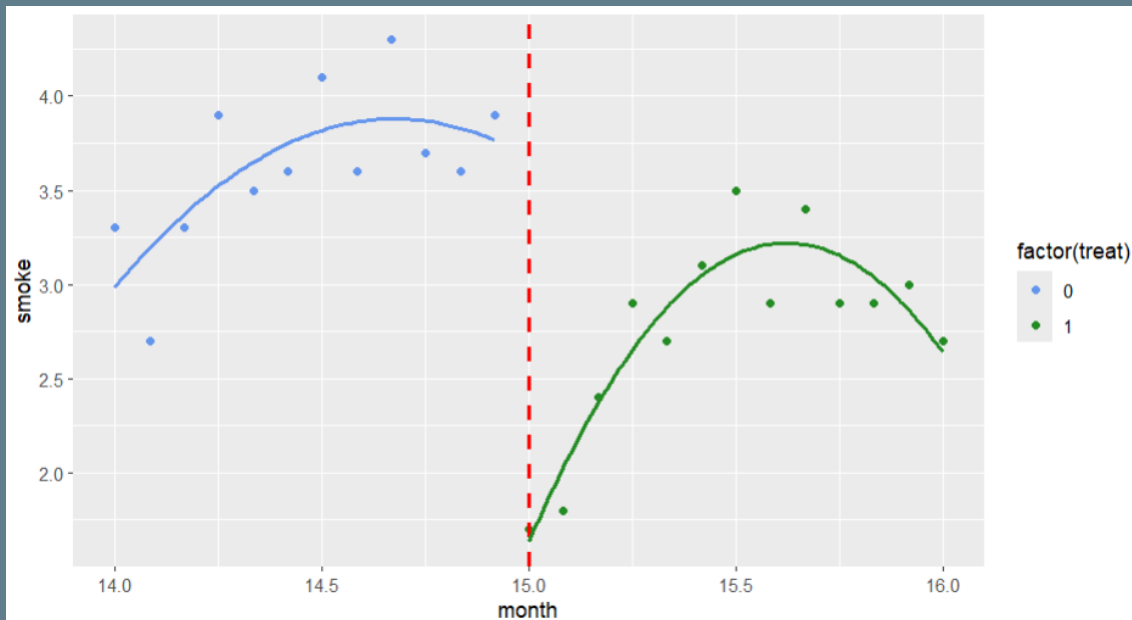
	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	4.0864	0.2448	16.690	1.35e-13	***
treat	-1.8259	0.3216	-5.678	1.23e-05	***
month_d	0.8517	0.3992	2.134	0.0448	*
treat:month_d	0.1504	0.5335	0.282	0.7807	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RDD 분석 - 대역폭 및 모형 선택

대역폭 12개월

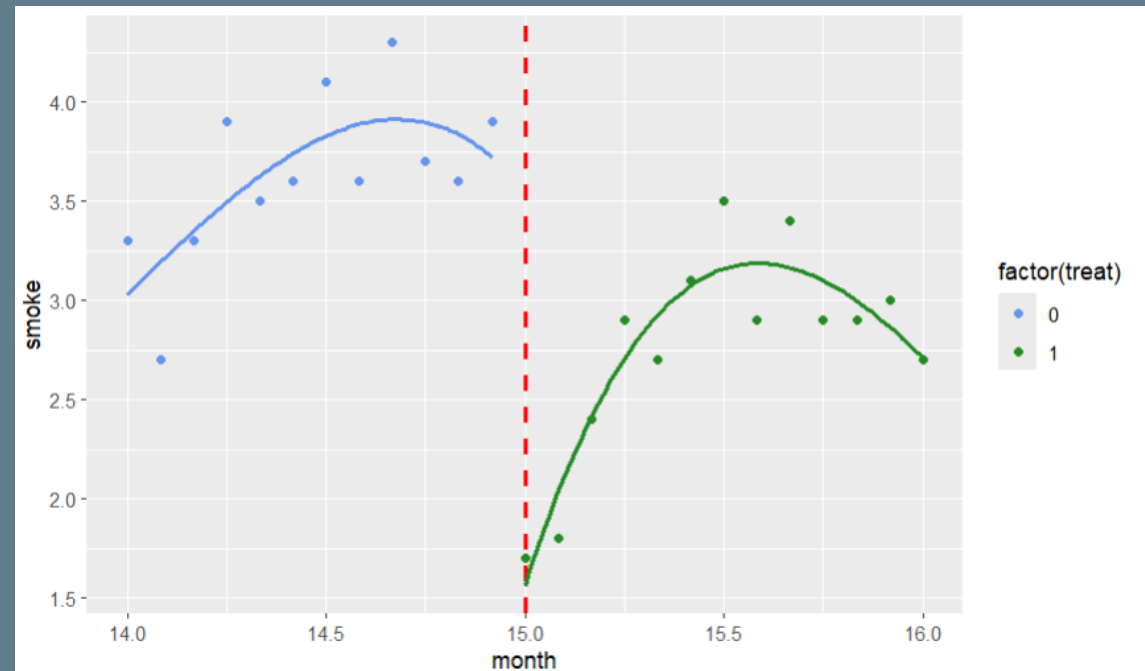
비선형 RDD 회귀모형(2차항) 결과 시각화 및 모수 추정 결과



Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	3.6750	0.2871	12.799	8.65e-11	***
treat	-2.0387	0.3497	-5.830	1.29e-05	***
month_d	-1.2638	1.2186	-1.037	0.312700	
I(month_d^2)	-1.9528	1.0950	-1.783	0.090509	.
treat:month_d	6.3515	1.5315	4.147	0.000547	***
treat:I(month_d^2)	-2.1327	1.4137	-1.509	0.147853	

비선형 RDD 회귀모형(3차항) 결과 시각화 및 모수 추정 결과



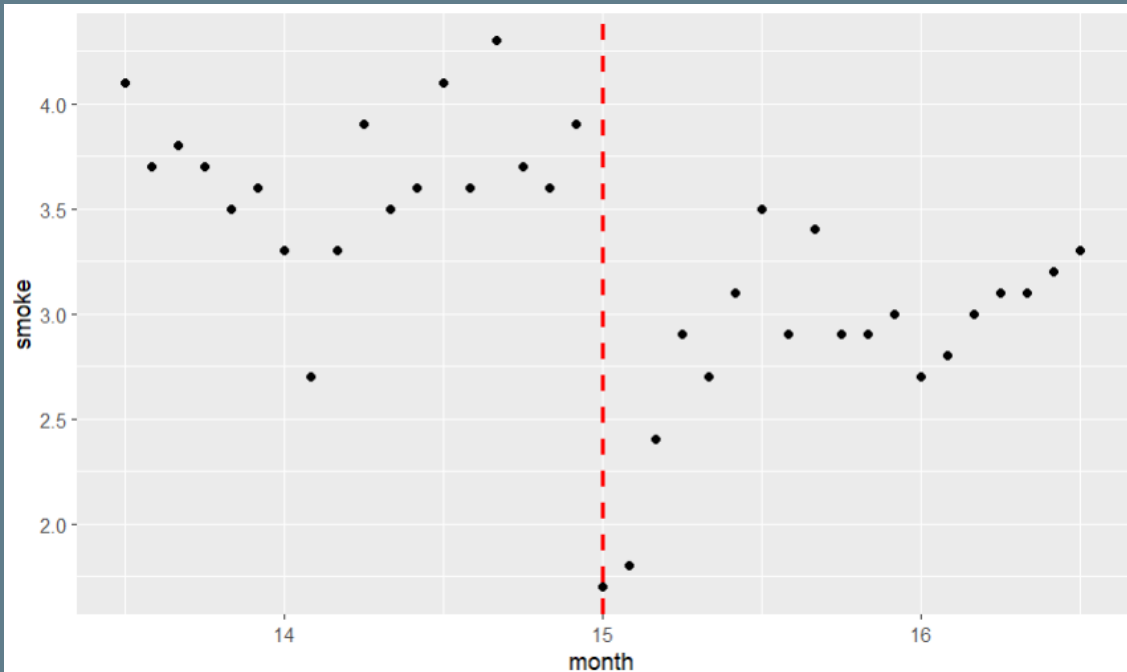
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	3.5566	0.4747	7.492	8.81e-07	***
treat	-1.9953	0.5354	-3.727	0.00168	**
month_d	-2.3623	3.6425	-0.649	0.52531	
I(month_d^2)	-4.3892	7.6559	-0.573	0.57394	
I(month_d^3)	-1.4993	4.6585	-0.322	0.75149	
treat:month_d	8.5818	4.2734	2.008	0.06078	.
treat:I(month_d^2)	-2.6418	9.3247	-0.283	0.78036	
treat:I(month_d^3)	3.4629	5.8231	0.595	0.55988	

RDD 분석 - 대역폭 및 모형 선택

대역폭 18개월

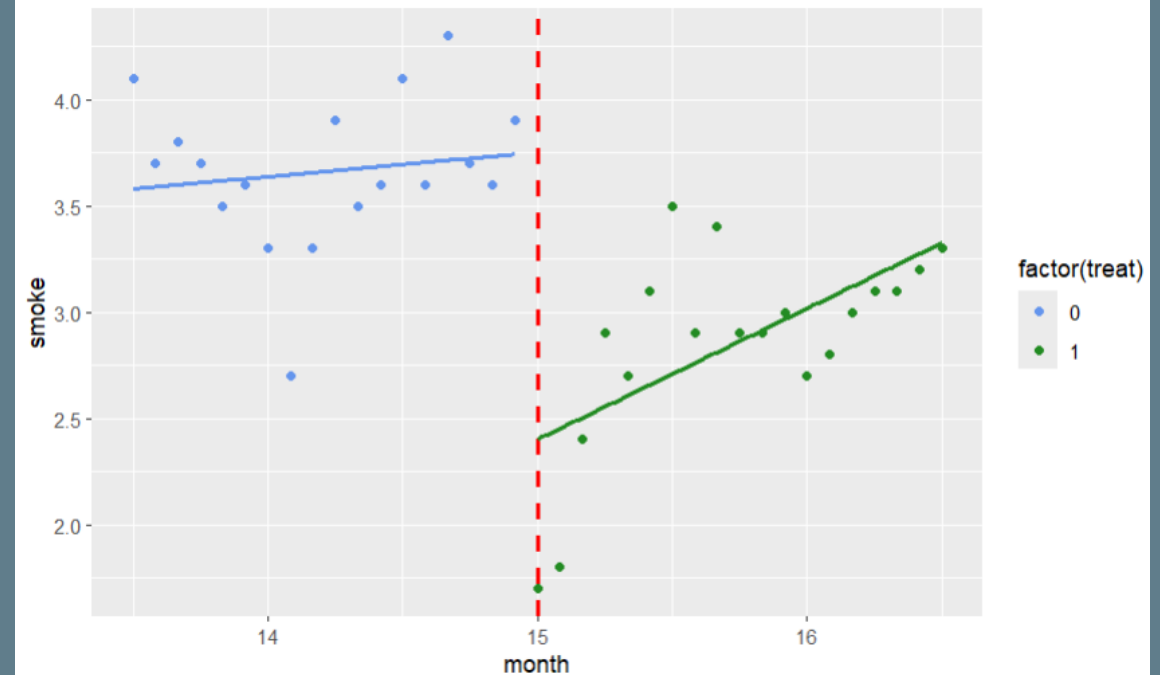
대역폭 6개월 기본 시각화



변수 목록

month: 년월 (개월 기준 소수점 표기)
increase: 담뭇값 인상 년월
month_d: 년월 - 담뭇값 인상 년월
treat: 담뭇값 인상 여부
smoke: 담배 판매량 (단위:억갑)

RDD 회귀모형 결과 시각화 및 모수 추정 결과



Coefficients:

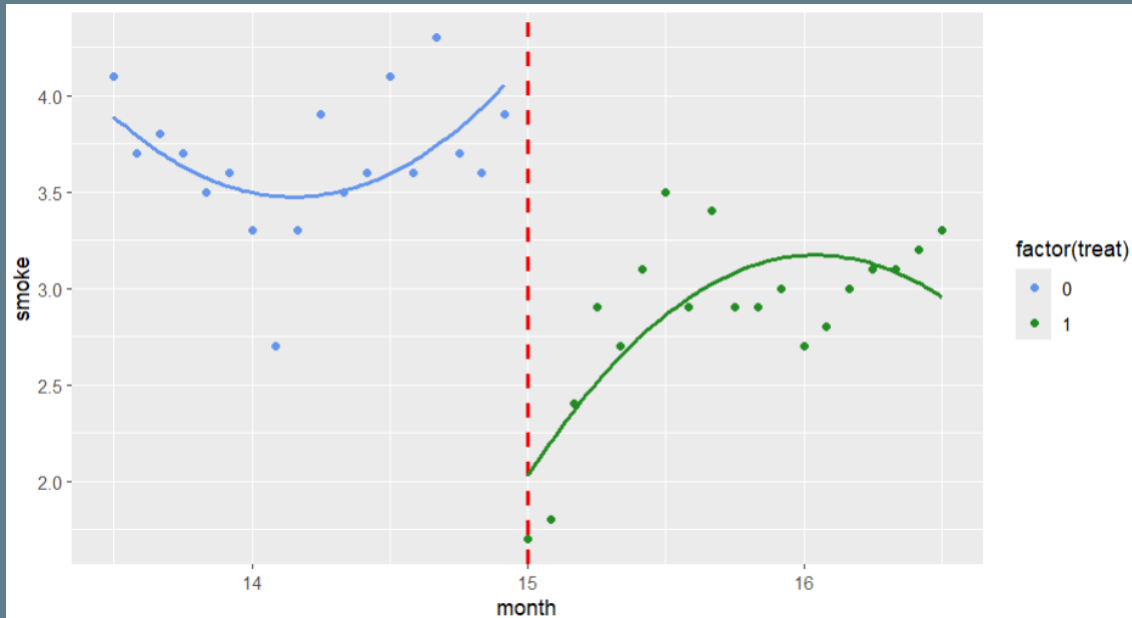
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.7523	0.1835	20.445	< 2e-16 ***
treat	-1.3518	0.2466	-5.482	4.45e-06 ***
month_d	0.1152	0.2035	0.566	0.575
treat:month_d	0.5017	0.2767	1.813	0.079 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

RDD 분석 - 대역폭 및 모형 선택

대역폭 18개월

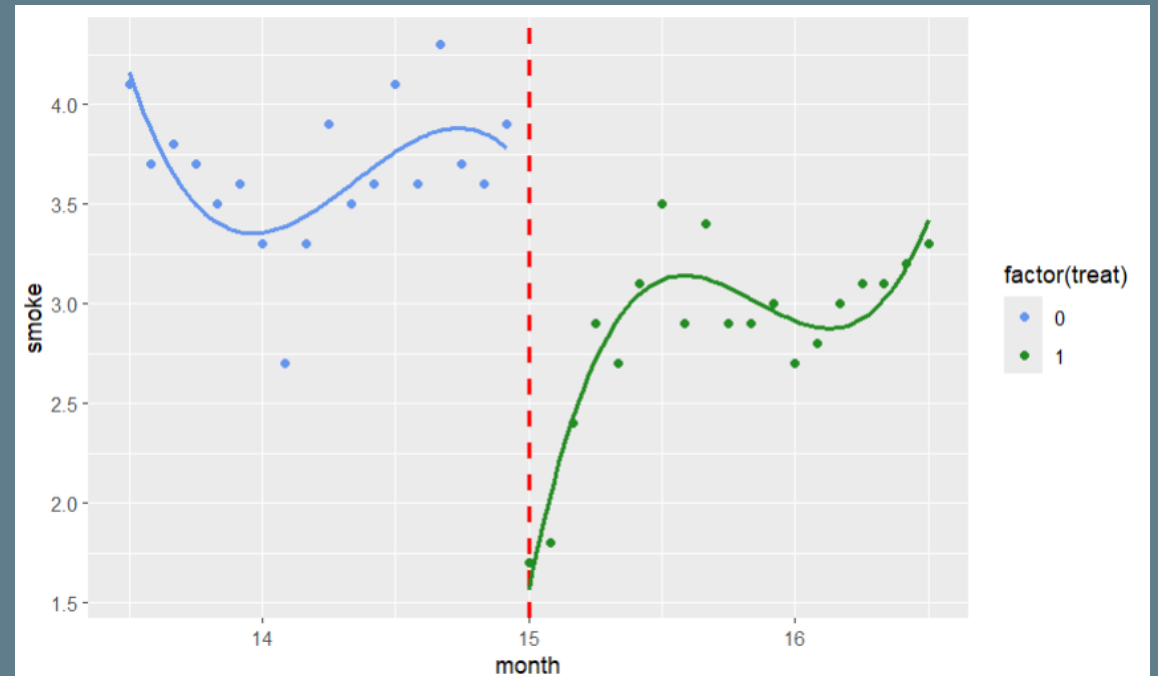
비선형 RDD 회귀모형(2차항) 결과 시각화 및 모수 추정 결과



Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.1853	0.2625	15.942	< 2e-16 ***
treat	-2.1566	0.3335	-6.467	3.28e-07 ***
month_d	1.6740	0.7634	2.193	0.03596 *
I(month_d^2)	0.9845	0.4685	2.101	0.04384 *
treat:month_d	0.5175	0.9934	0.521	0.60607
treat:I(month_d^2)	-2.0343	0.6219	-3.271	0.00263 **

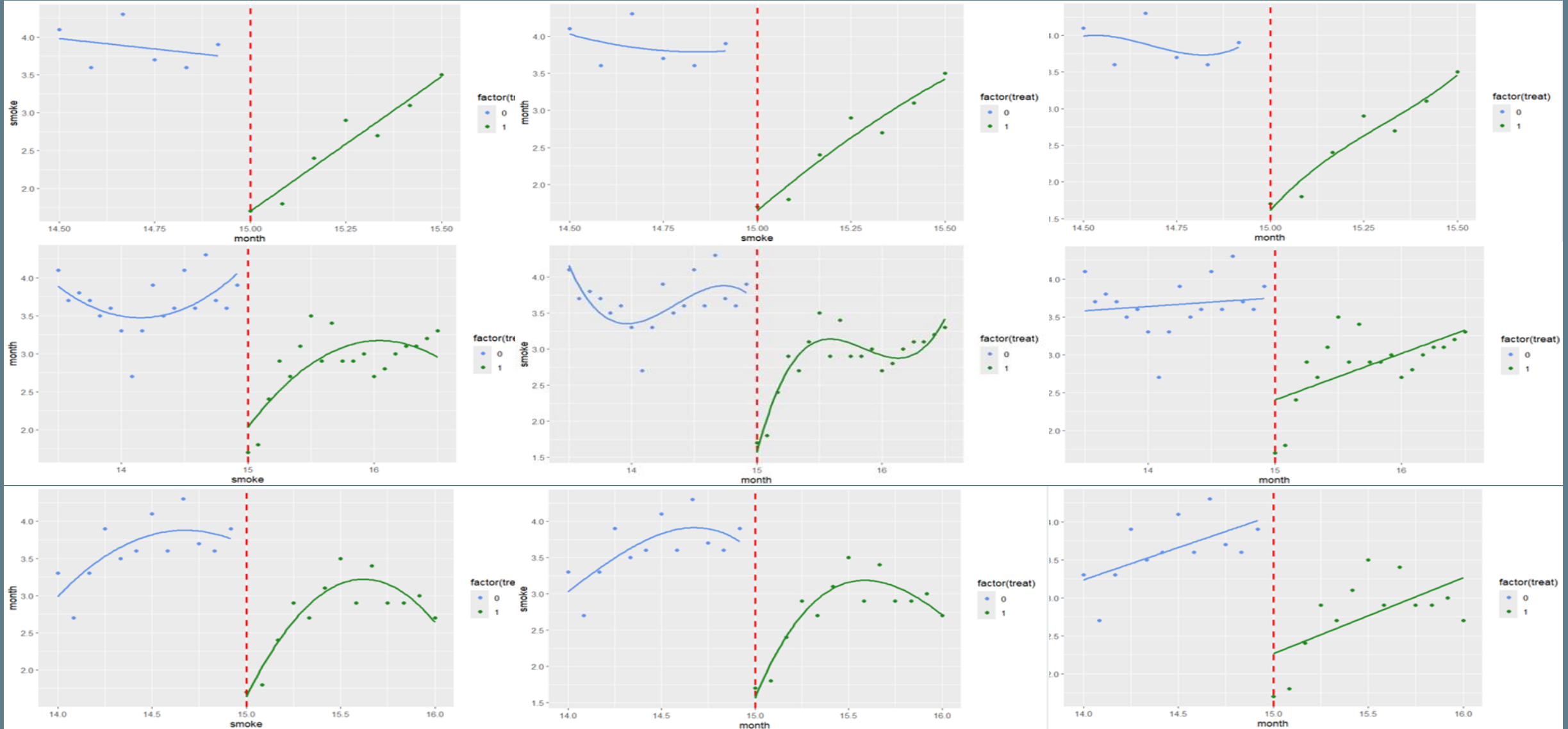
비선형 RDD 회귀모형(3차항) 결과 시각화 및 모수 추정 결과



Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.6534	0.3016	12.113	7.21e-13 ***
treat	-2.0929	0.3592	-5.827	2.56e-06 ***
month_d	-1.8851	1.6052	-1.174	0.249796
I(month_d^2)	-4.4861	2.3228	-1.931	0.063271 .
I(month_d^3)	-2.3034	0.9661	-2.384	0.023871 *
treat:month_d	8.4192	1.9790	4.254	0.000200 ***
treat:I(month_d^2)	-4.0003	2.9524	-1.355	0.185902
treat:I(month_d^3)	5.6086	1.2528	4.477	0.000108 ***

결론 - 최종 대역폭 및 모형 선택

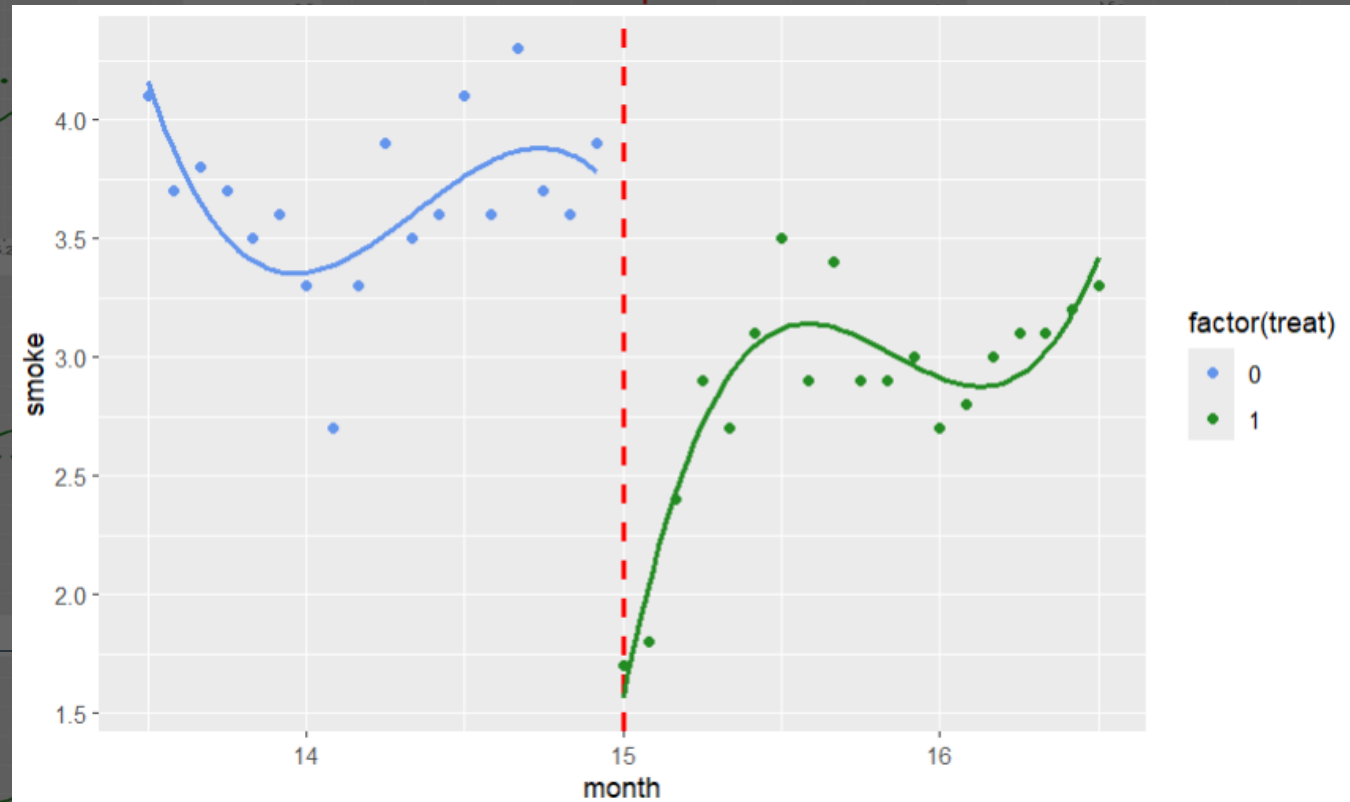


결론 - 최종 대역폭 및 모형 선택



결론 - 최종 대역폭 및 모형 선택

1년 6개월 / 비선형 RDD모형 (3차항)



추정된 LATE 값 : -2.0929 / 임계값 : 2015년 1월

담뱃값 인상이라는 처치가 담배 판매량(흡연율)에 감소하는 방향으로 유의미한 영향을 미친 것으로 나타남

RDD 결론

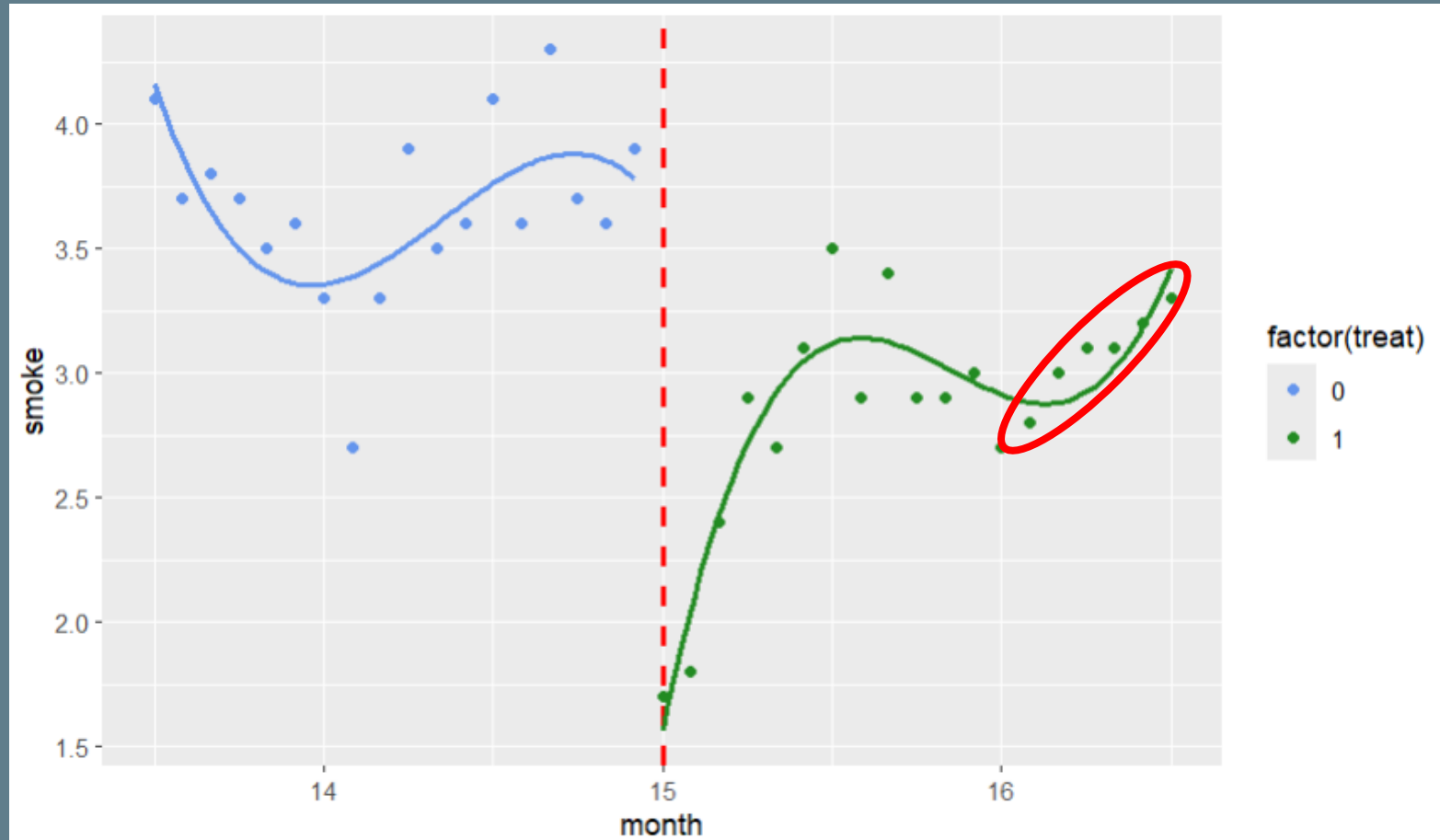
RDD 분석 결과 :

담뱃값 인상이 담배 판매량(흡연율)에 단기적으로 유의미한 영향을 미친 것으로 나타남

→ 담뱃값의 가격 변화가 담배 판매량(흡연율)에 중요한 역할을 한다는 것을 시사함

→ 가격 정책이 흡연율을 줄이기 위한 효과적인 수단이 될 수 있음을 보여줌

하지만...



단기적 감소 이후 다시 원상복구 되는 것처럼 증가하는 추세를 보이는데...?

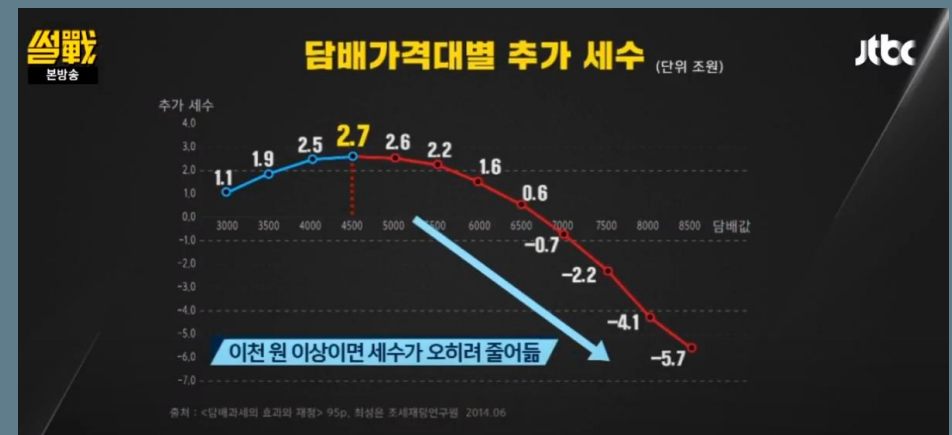
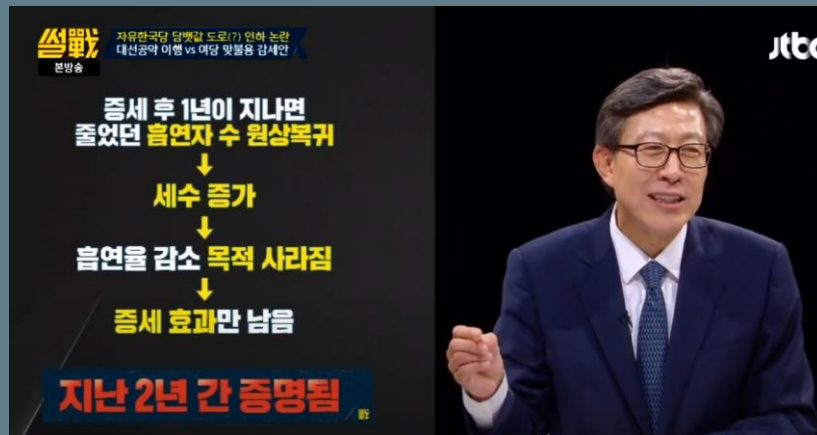
인과성 분석(DID) 배경

RDD 분석 결과 :

담뱃값 인상이 담배 판매량(흡연율)에 단기적으로 유의미한 영향을 미친 것으로 나타남

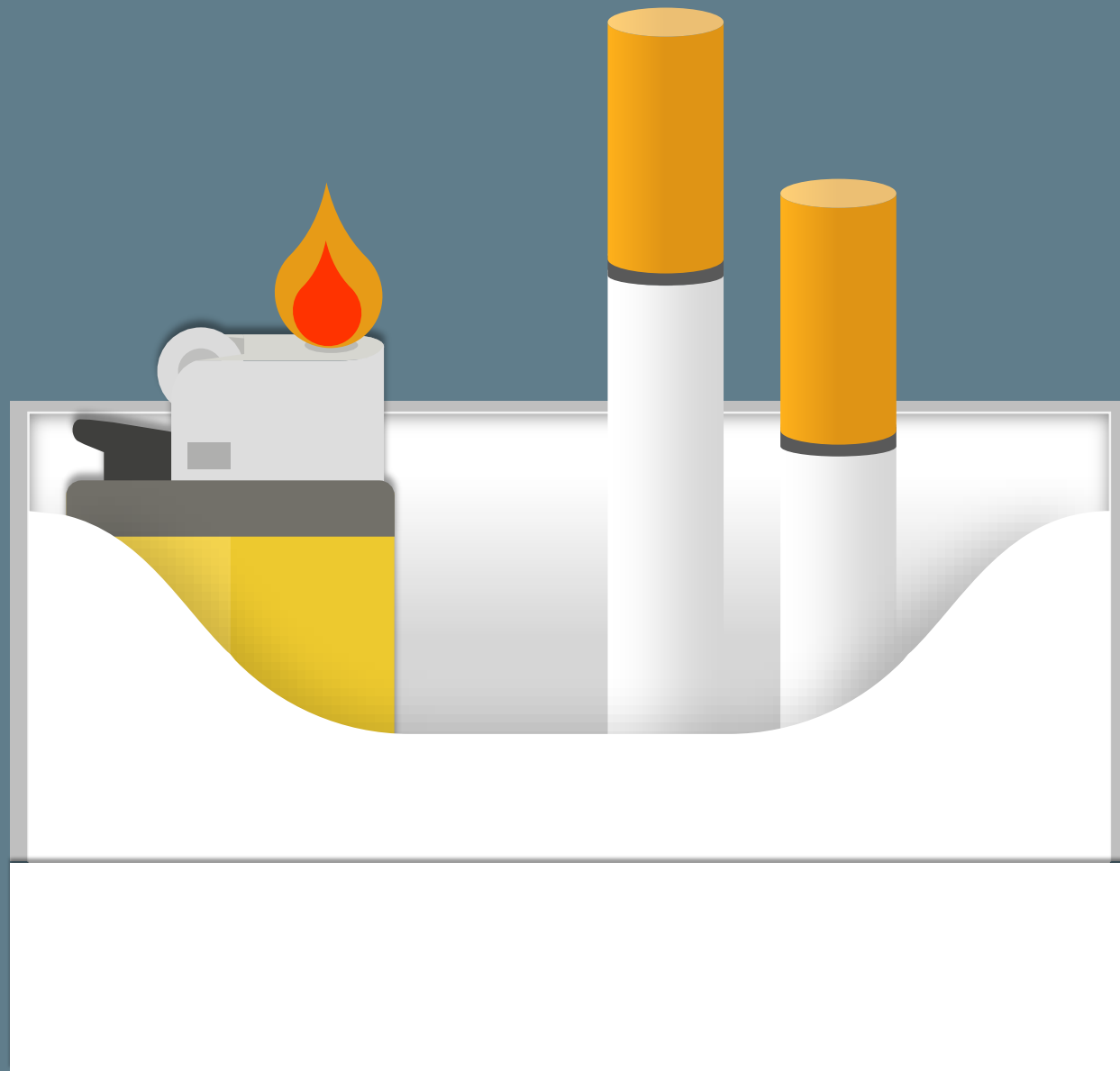
DID 배경 :

하지만, 2015년 담뱃값 인상은 과연 세수가 목적이 아닌
장기적으로도 흡연율이 감소한 정책이었는가?



인과성 분석

DID



인과성 분석(DID) 개요 및 데이터 수집

분석 목적 : 2015년 담뭇값 인상이 한국 성인 흡연에 미친 장기적 영향 분석

처치:

2015년 담뭇값 인상

처치군:

한국의 매일 흡연율 데이터

통제군:

일본의 매일 흡연율 데이터

통제 변수:

국가, 연도, GDP

수집 기간:

2011년 ~ 2018년

연간 매일 흡연율

출처:

OECD Statistics 2023

DID 가정 성립 확인

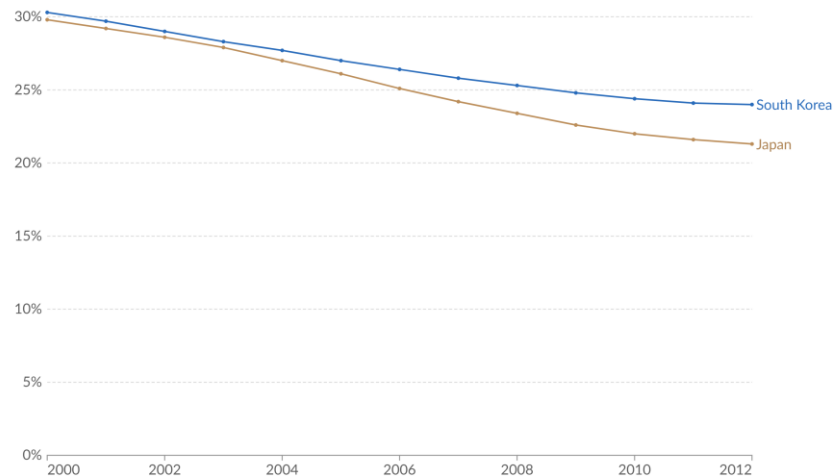
병행추세 검정

사용 방법: 결과변수들의 처치 이전 시점 추이 **시각적** 비교, FCTC 가입

사용된 변수: 매일 흡연율, GDP, 성인 흡연율

Share of people who smoke every day, 2000 to 2012

Estimates of the prevalence of daily smoking, defined as the percentage of men and women, of all ages, who smoke daily.



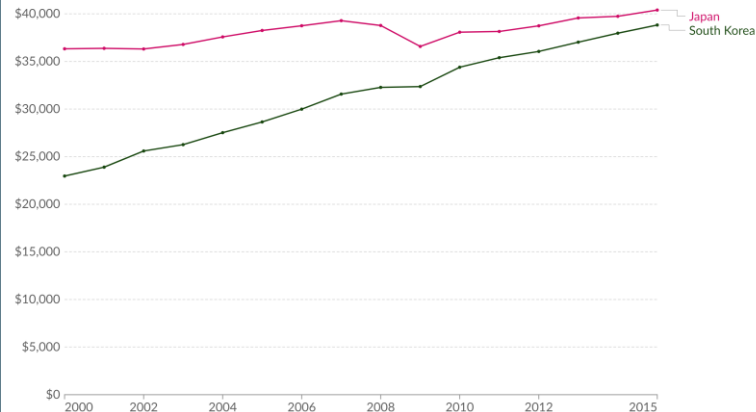
Data source: Institute for Health Metrics and Evaluation (IHME) (2012)

OurWorldInData.org/smoking | CC BY

2000~2012년의 매일 흡연율 변화 추이

GDP per capita

This data is adjusted for inflation and for differences in the cost of living between countries.



Data source: World Bank (2023)

Note: This data is expressed in international-\$* at 2017 prices.

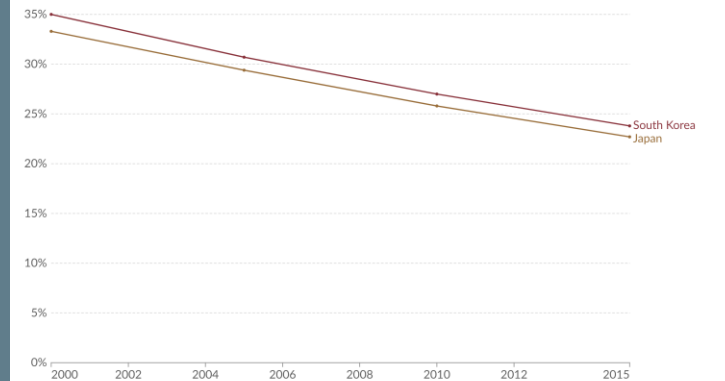
OurWorldInData.org/economic-growth | CC BY

1. **International dollars:** International dollars are a hypothetical currency that is used to make meaningful comparisons of monetary indicators of living standards. Figures expressed in international dollars are adjusted for inflation within countries over time, and for differences in the cost of living between countries. The goal of such adjustments is to provide a unit whose purchasing power is held fixed over time and across countries, such that one international dollar can buy the same quantity and quality of goods and services no matter where or when it is spent. Read more in our article: What are Purchasing Power Parity adjustments and why do we need them?

2000~2015년의 GDP 변화 추이

Share of adults who smoke, 2000 to 2015

Share of people aged 15 and older who smoke any tobacco product on a daily or non-daily basis. It excludes smokeless tobacco use. Smoking is a risk factor* for chronic complications, including cancers, cardiovascular disease*, and premature death.



Data source: Multiple sources compiled by World Bank (2024)

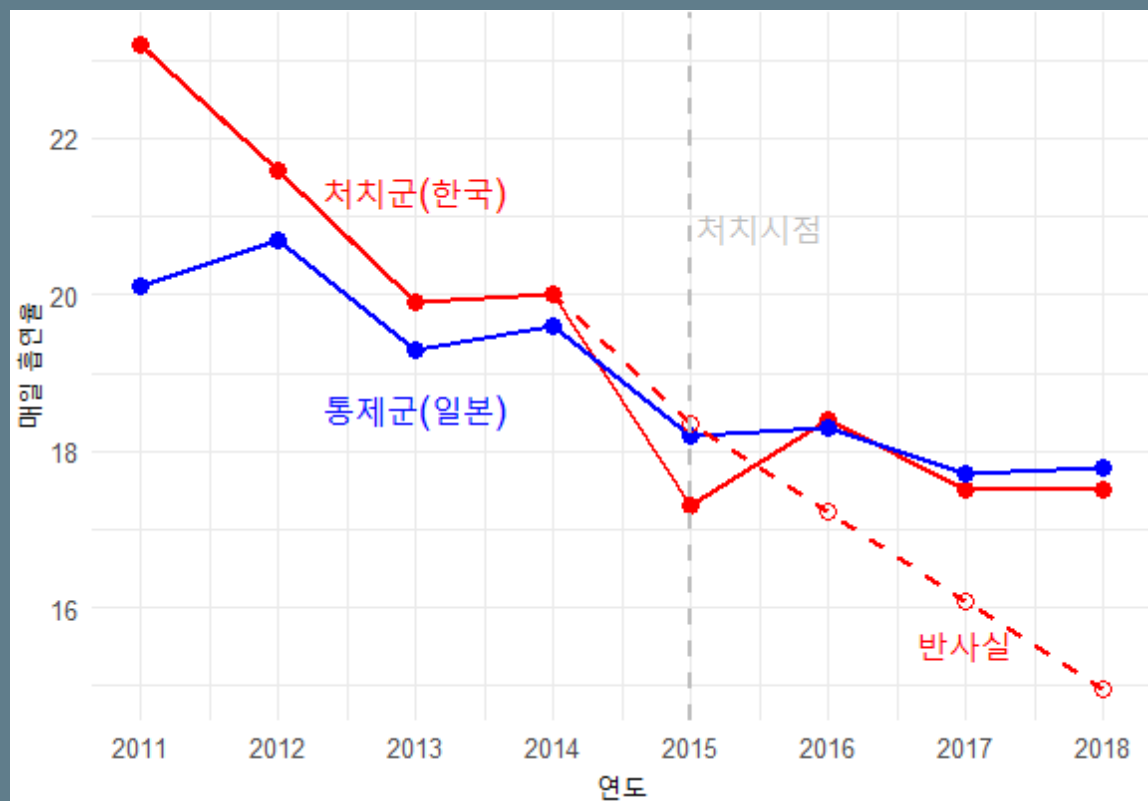
OurWorldInData.org/smoking | CC BY

1. **Risk factor:** A risk factor is a condition or behavior that increases the likelihood of developing a given disease or injury, or an outcome such as death. The impact of a risk factor is estimated in different ways. For example, a common approach is to estimate the number of deaths that would occur if the risk factor was absent. Risk factors are not mutually exclusive: people can be exposed to multiple risk factors, which contribute to their disease or death. Because of this, the number of deaths caused by each risk factor is typically estimated separately. Read more: How do researchers estimate the death toll caused by each risk factor, whether it's smoking, obesity or air pollution? Read more: Why isn't it possible to sum up the death toll from different risk factors?

2. **Cardiovascular disease:** Cardiovascular diseases cover all diseases of the heart and blood vessels - including heart attacks and strokes, atherosclerosis, ischemic heart disease, hypertensive diseases, cardiomyopathy, rheumatic heart disease, and more. They tend to develop gradually with age, especially when people have risk factors like high blood pressure, smoking, alcohol use, poor diet, and air pollution.

2000~2015년의 성인 흡연율 변화 추이

DID 분석 결과



```
> summary(did_model)
```

Call:

```
lm(formula = 흡연율 ~ 국가 + 연도 + 처치후 + GDP, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.14727	-0.42148	0.03716	0.41866	1.29816

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.493e+02	7.431e+02	1.008	0.335
국가	9.563e-01	1.462e+00	0.654	0.526
연도	-3.595e-01	3.804e-01	-0.945	0.365
처치후	-1.487e+00	1.037e+00	-1.434	0.179
GDP	-1.541e-04	6.039e-04	-0.255	0.803

Residual standard error: 0.7454 on 11 degrees of freedom
Multiple R-squared: 0.8551, Adjusted R-squared: 0.8024
F-statistic: 16.22 on 4 and 11 DF, p-value: 0.0001389

DID 분석 결과

```
> summary(did_model)

call:
lm(formula = 흡연율 ~ 국가 + 연도 + 처치후 + GDP, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.14727 -0.42148  0.03716  0.41866  1.29816

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.493e+02  7.431e+02   1.008   0.335
국가          9.563e-01  1.462e+00   0.654   0.526
연도        -3.595e-01  3.804e-01  -0.945   0.365
처치후       -1.487e+00  1.037e+00  -1.434   0.179
GDP         -1.541e-04  6.039e-04  -0.255   0.803

Residual standard error: 0.7454 on 11 degrees of freedom
Multiple R-squared:  0.8551,    Adjusted R-squared:  0.8024
F-statistic: 16.22 on 4 and 11 DF,  p-value: 0.0001389
```

1) 국가나 연도, GDP에 따른 흡연율 차이

p-value > 0.05 → 유의미 X

2) 처치 효과 검정

중요한 것은 처치 후 변수

DID Estimator < 0 이지만 그 값이 매우 작음

p-value가 유의 X

→ 처치가 흡연율 감소를 일으켰을 수도 있지만,
장기적 처치 효과 통계적 유의 X

3) 한계점

데이터 포인트의 갯수 매우 적은 편

→ 처치 효과 유의미하게 포착하지 못했을 수 있음

DID 결론

DID 분석 결과 :

2015년의 처치는 단기적으로는 효과가 있을지 몰라도
장기적 처치 효과는 통계적으로 유의하지 않았음을 확인할 수 있음

차후 유의미한 흡연율 감소를 발생시키려면 최대 세수보다는 조금 더 담뱃값을 인상시켜
흡연자 수가 원상복구되지 않을 정도의 인상이 필요함

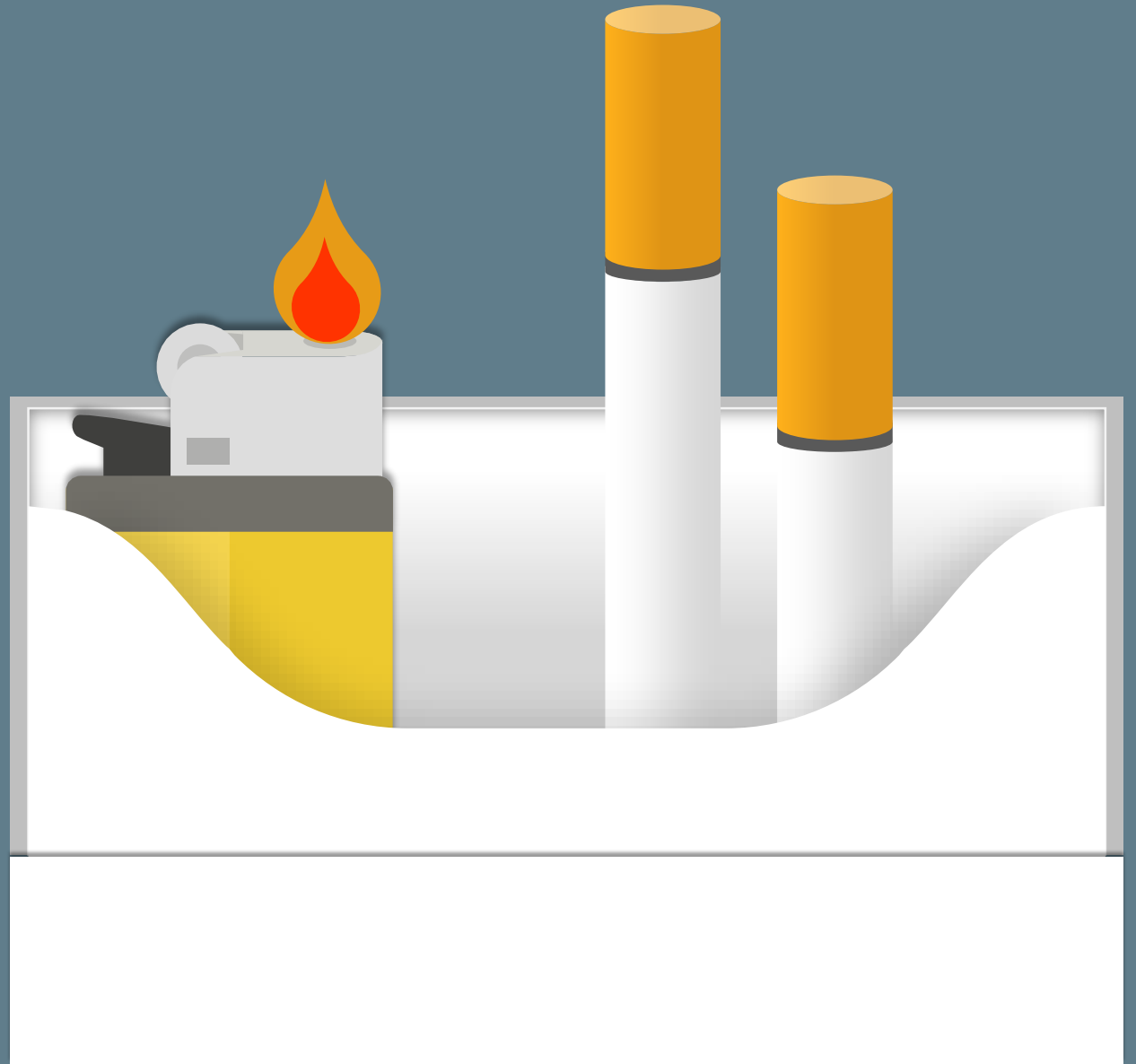
그렇다면...

도입부처럼 올해 담뱃값 인상이 발생했다면
월간 담배 판매량(흡연율)은 어떻게 되었을까?

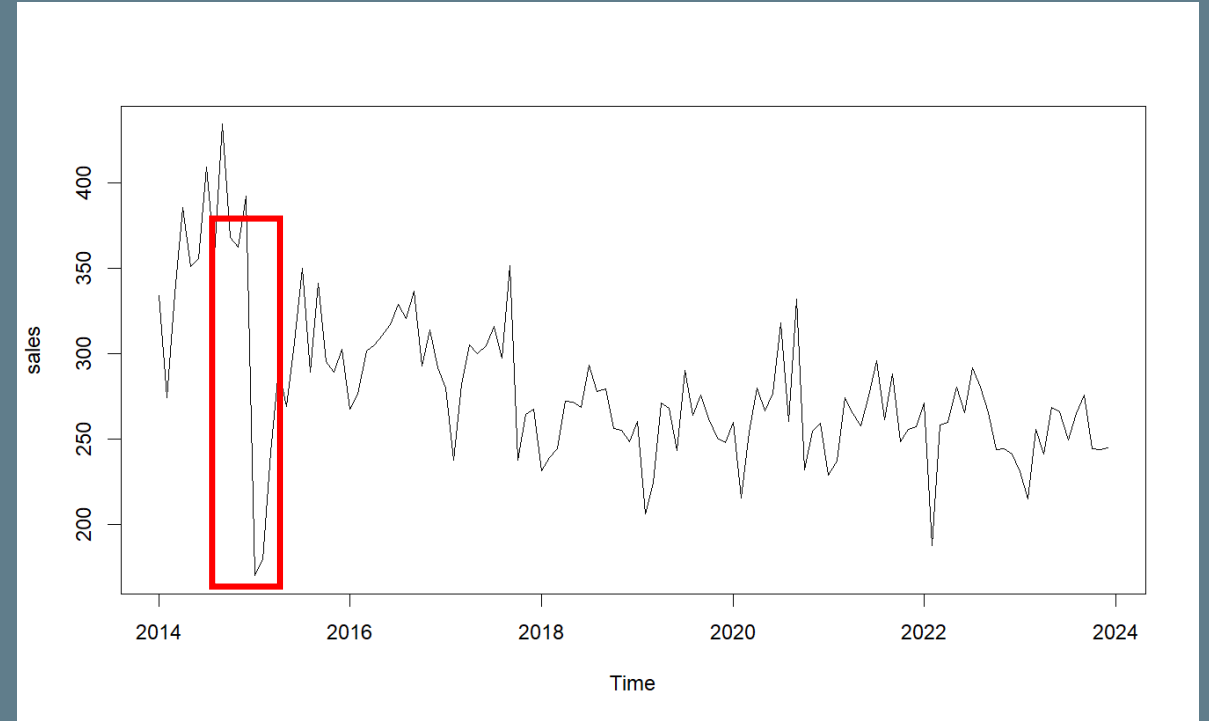
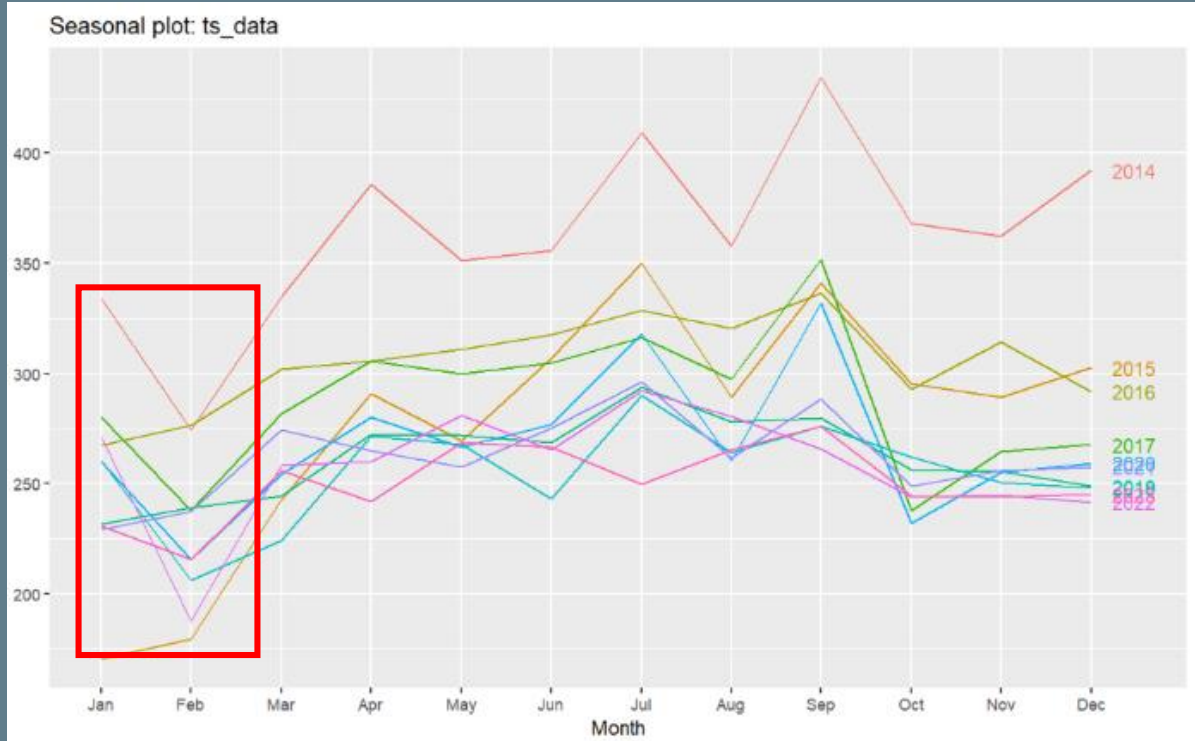
앞에서 RDD, DID로 분석 결과에서 도출된 처치 효과 기반
담뱃값 인상이 월간 담배 판매량(흡연율)에 미치는 영향 평가

위의 분석 결과와 시계열 분석을 통해 올해 월간 담배 판매량을 예측
→ 담뱃값 인상이 소비자 행동과 시장에 미치는 잠재적 영향 파악

시 계 열 분 석
SARIMA

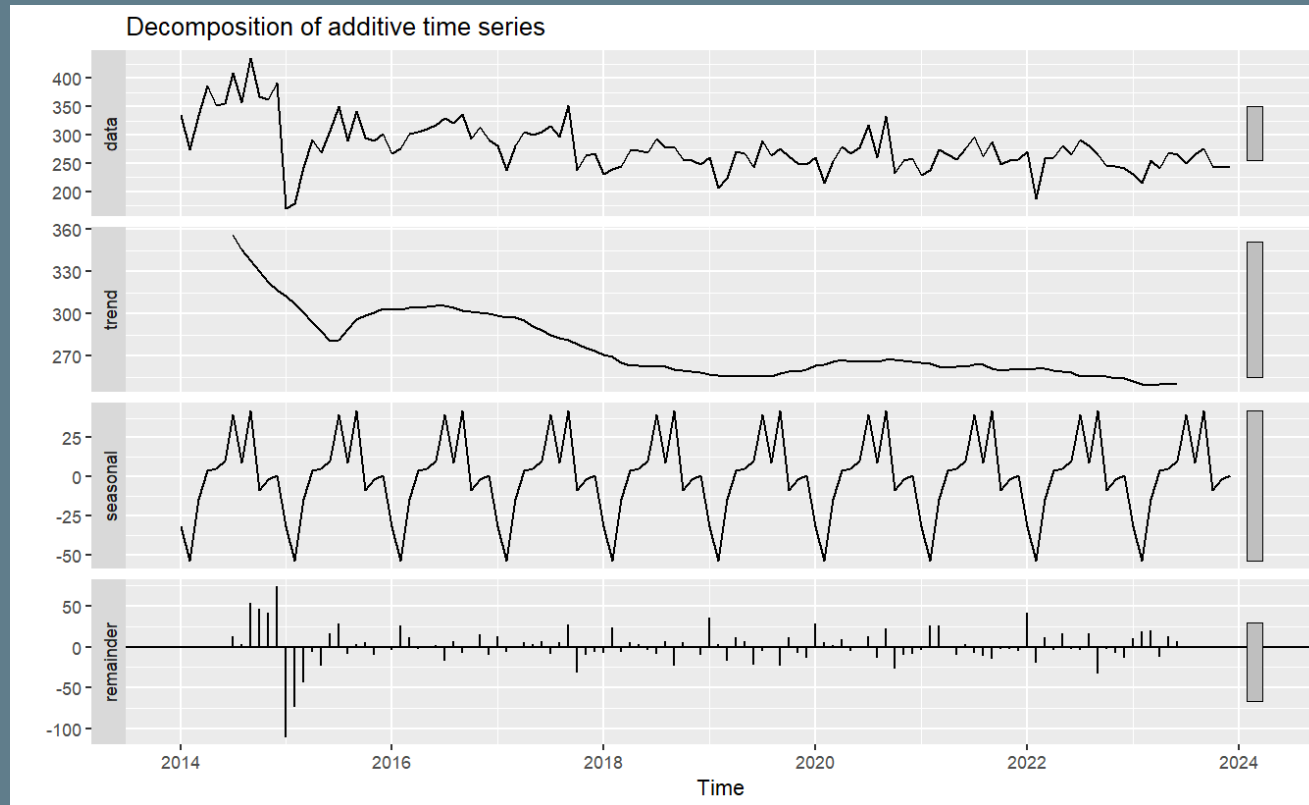


분석 준비 - 데이터 파악



1. 신년 다짐으로 연초에 담배 판매량이 하락하는 것으로 보임.
2. 2015년 담배값 인상 당시, 담배 판매량이 급락한 것을 확인할 수 있음.

시계열 성분 분해



1년 주기의 뚜렷한 계절성을 확인

2015년 담배값 4500원 인상 당시

급격한 감소 이후, 약간의 증가

→ 전반적으로 감소하는 추세까지 확인

비교적 일정한 변동성 → 가법 모형 선택

단위근 검정(시계열 정상성) - DF검정

```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
  
Test regression trend  
  
Call:  
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)  
  
Residuals:  
      Min       1Q   Median       3Q      Max   
-178.913  -14.315    1.737   16.566   89.277  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)      
(Intercept) 155.13610    32.15650   4.824 4.39e-06 ***  
z.lag.1      -0.48312     0.09836  -4.912 3.04e-06 ***  
tt           -0.34602     0.11294  -3.064 0.00273 **  
z.diff.lag   -0.15479     0.09155  -1.691 0.09361 .  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 33.41 on 114 degrees of freedom  
Multiple R-squared:  0.3068,    Adjusted R-squared:  0.2886  
F-statistic: 16.82 on 3 and 114 DF,  p-value: 4.11e-09  
  
Value of test-statistic is: -4.9117 8.059 12.0725  
  
Critical values for test statistics:  
      1pct   5pct  10pct  
tau3  -3.99  -3.43  -3.13  
phi2   6.22   4.75   4.07  
phi3   8.43   6.49   5.47
```

1) DF 통계량 $r = -6.8618$

→ 절대값 기준 1% 유의 수준 임계값 보다 큼 → 단위근 존재X

→ 정상 시계열

2) drift 통계량 = 8.069

→ 절대값 기준 1% 유의 수준 임계값보다 큼 → 단위근 존재X

→ 정상 시계열

3) trend = 23.5472

→ 절대값 기준 1% 유의 수준 임계값보다 큼 → 단위근 존재X

→ 정상 시계열

추세 + 상수항을 가정했을 때 DF 검정 결과,
월별 담배 판매량 데이터는 정상 시계열 데이터임

단위근 검정(시계열 정상성) - ADF검정

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-178.913  -14.315    1.737   16.566   89.277

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  155.13610   32.15650   4.824 4.39e-06 ***
z.lag.1       -0.48312    0.09836  -4.912 3.04e-06 ***
tt            -0.34602    0.11294  -3.064 0.00273 **
z.diff.lag    -0.15479    0.09155  -1.691 0.09361 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 33.41 on 114 degrees of freedom
Multiple R-squared:  0.3068,    Adjusted R-squared:  0.2886
F-statistic: 16.82 on 3 and 114 DF,  p-value: 4.11e-09

Value of test-statistic is: -4.9117 8.059 12.0725

Critical values for test statistics:
      1pct  5pct 10pct
tau3  -3.99 -3.43 -3.13
phi2   6.22  4.75  4.07
phi3   8.43  6.49  5.47
```

1) DF 통계량 $r = -4.9117$

→ 절대값 기준 1% 유의 수준 임계값 보다 큼 → 단위근 존재X

→ 정상 시계열

2) drift 통계량 = 15.7175

→ 절대값 기준 1% 유의 수준 임계값보다 큼 → 단위근 존재X

→ 정상 시계열

3) trend = 12.0725

→ 절대값 기준 1% 유의 수준 임계값보다 큼 → 단위근 존재X

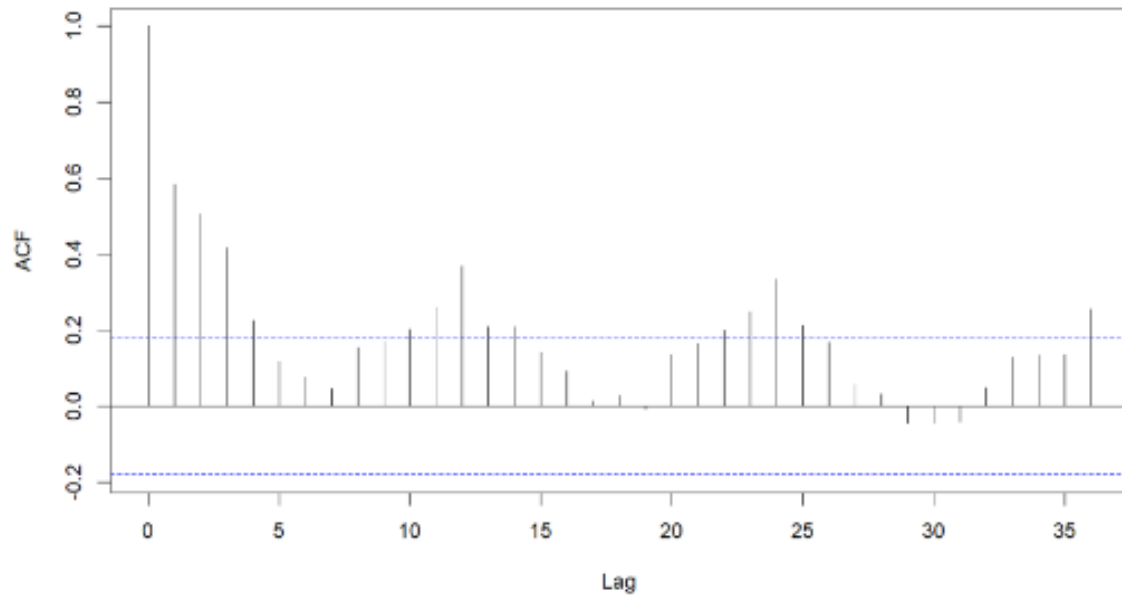
→ 정상 시계열

추세 + 상수항을 가정했을 때 ADF 검정 결과,
월별 담배 판매량 데이터는 정상 시계열 데이터임

ACF, PACF를 통한 AR, MA 모수 추정

6.1 원본 시계열 데이터 ACF 확인

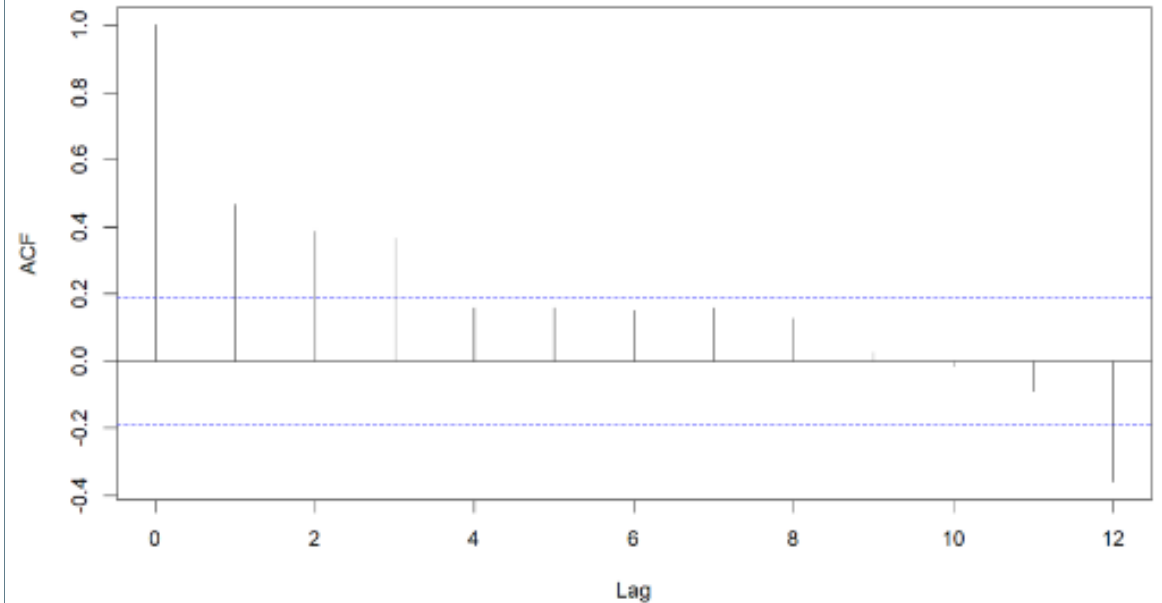
origin ACF Plot



뚜렷한 자기 상관 계절성이 확인됨
→ 계절성 차분 수행(12)

6.2 계절성 차분 이후, ACF 관찰

Seasonal ACF Plot

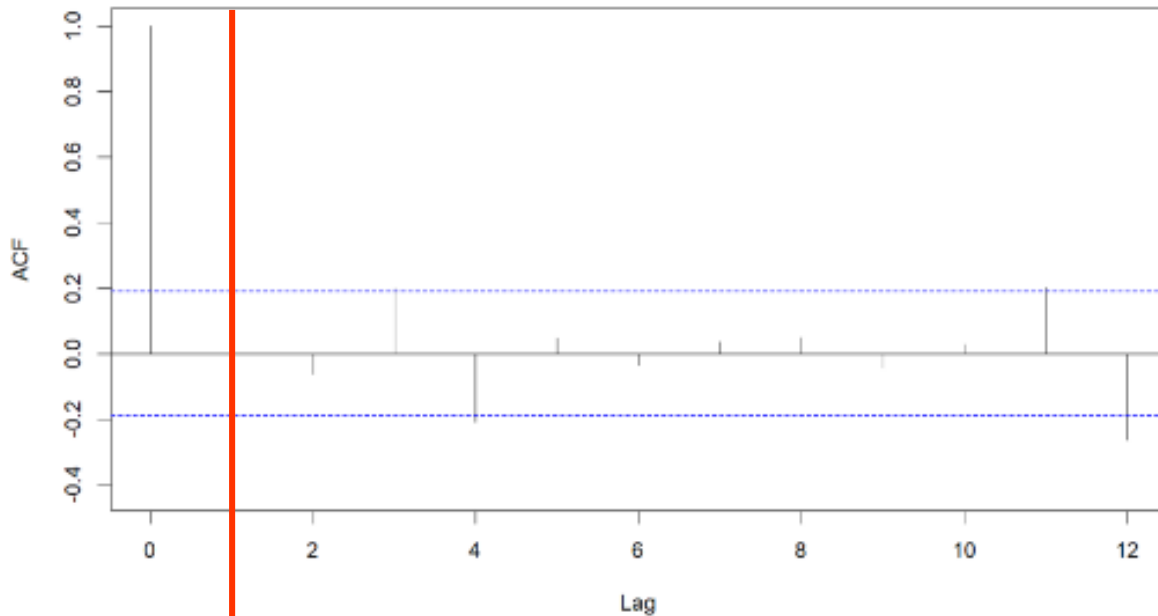


ACF 해석 결과 천천히 감소하는 패턴
→ 비정상성 데이터로 판단
→ 1차 차분 수행 ($d=1$)

ACF, PACF를 통한 AR, MA 모수 추정

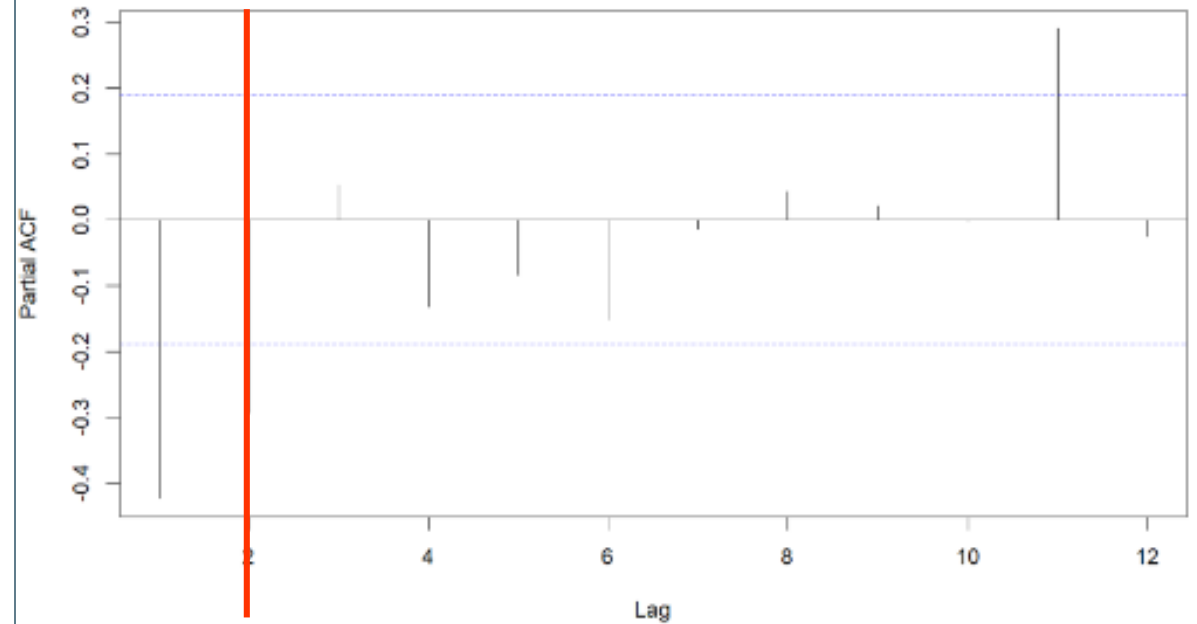
6.3 계절성 차분 + 일반 차분 이후, ACF, PACF 관찰

final ACF Plot



ACF plot에서 lag 1 이후 ACF 값이 급격히 감소
→ 정상성 확보 + $p = 0$

final PACF Plot



PACF plot에서 lag 2 이후 PACF 값이 급격히 감소
→ $q=1$

SARIMA 모형 설계

```
ARIMA(0,1,1)(0,1,0)[12]
```

```
Coefficients:
```

```
      ma1
```

```
    -0.6047
```

```
s.e.    0.0837
```

```
sigma^2 = 1017:  log likelihood = -521.99
```

```
AIC=1047.98   AICc=1048.1   BIC=1053.33
```

```
Training set error measures:
```

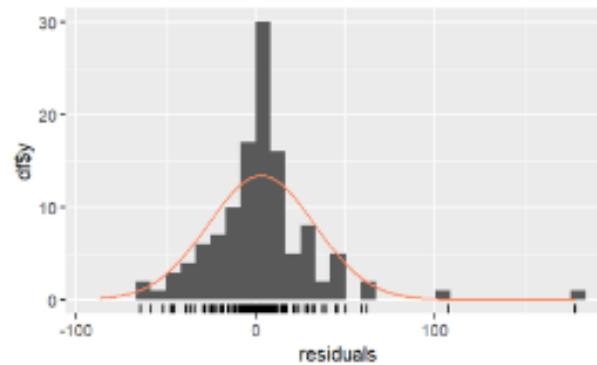
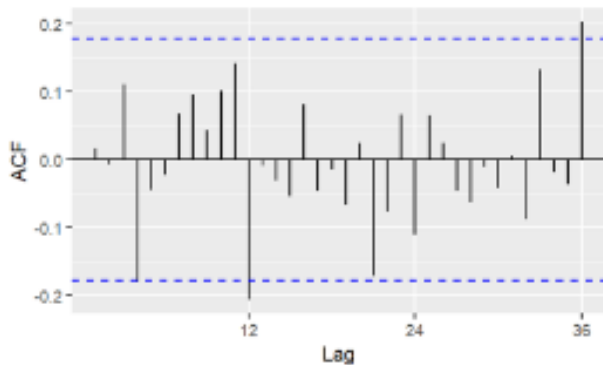
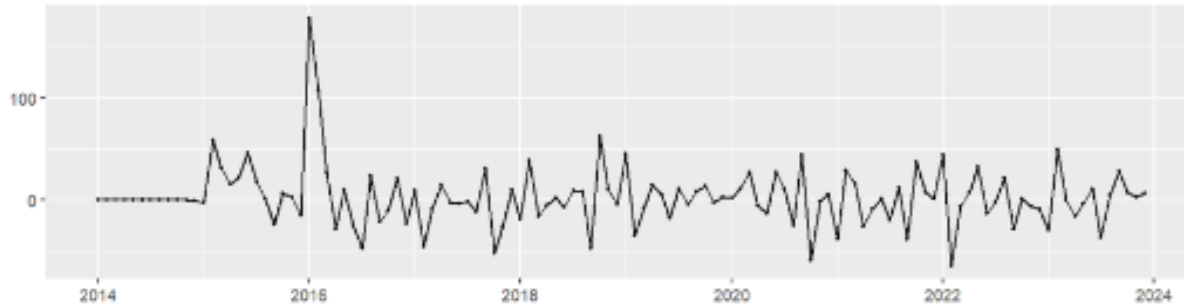
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	3.051826	29.96557	18.9863	1.03184	7.28222	0.6811909	0.01584903

비계절성 모수 : 앞에서 구한 값으로 설정 ($p=0, d=1, q=1$)

계절성 모수 : 차분만 설정 ($D=1$)

SARIMA 모형 설계

Residuals from ARIMA(0,1,1)(0,1,0)[12]



Ljung-Box test

data: Residuals from ARIMA(0,1,1)(0,1,0)[12]
 $Q^* = 28.335$, $df = 23$, **p-value = 0.2034**

Model df: 1. Total lags used: 24

모델 잔차의 백색잡음 검증
→ 잔차간의 자기상관 x

arima.auto를 활용한 자동 모델 설계

```
ARIMA(0,1,1)(0,1,1)[12]
```

Coefficients:

	ma1	ma2
	-0.6427	-0.6117
s.e.	0.0883	0.1098

```
sigma^2 = 792.8: log likelihood = -511.04
```

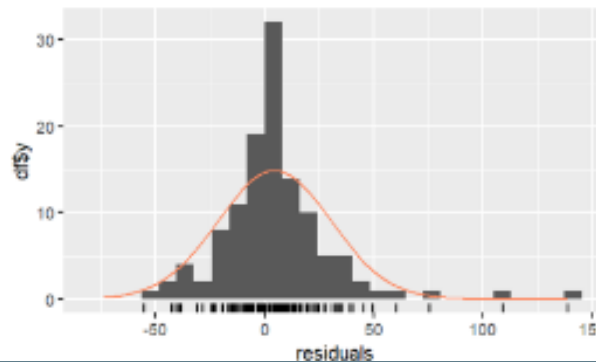
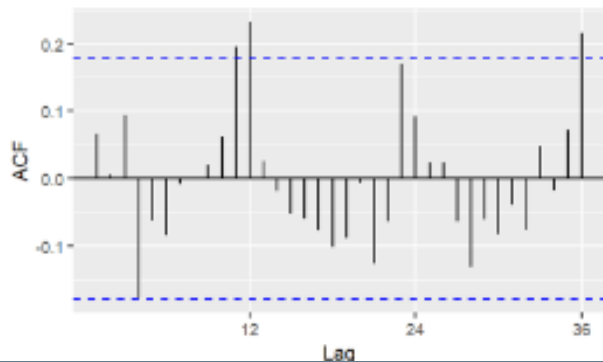
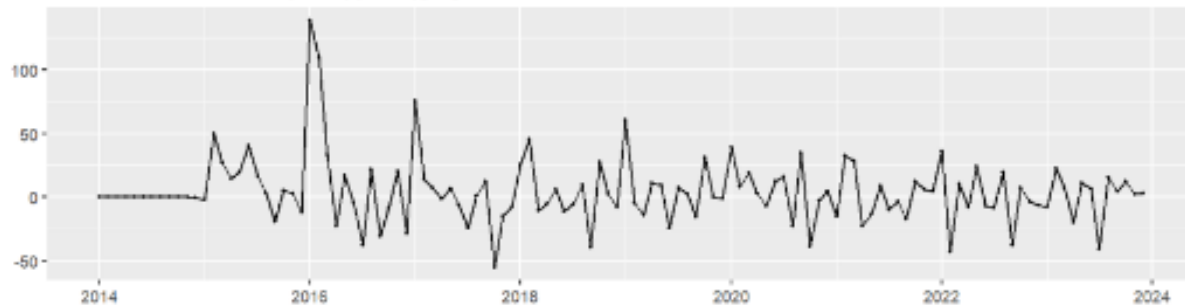
```
AIC=1028.07 AICc=1028.31 BIC=1036.09
```

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	4.498277	26.33781	16.945	1.662671	6.418097	0.607953	0.06636034

arima.auto를 활용한 자동 모델 설계

Residuals from ARIMA(0,1,1)(0,1,1)[12]



Ljung-Box test

data: Residuals from ARIMA(0,1,1)(0,1,1)[12]
Q* = 33.07, df = 22, **p-value = 0.06089**

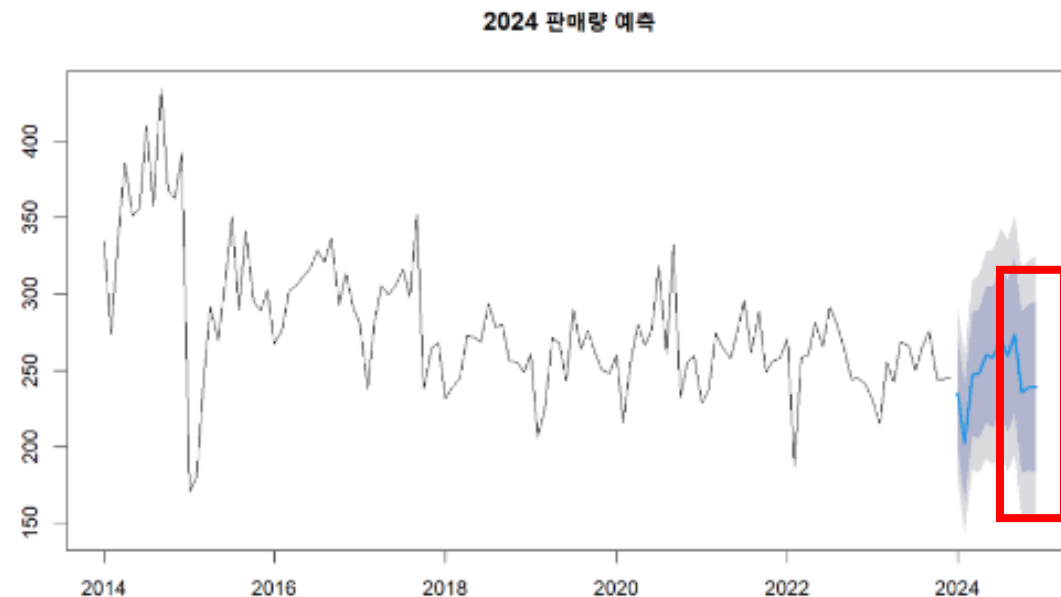
Model df: 2. Total lags used: 24

모델 잔차의 백색잡음 검증
→ 잔차간의 자기상관 x

2024 담배 판매량 예측

auto.arima로 학습된 모델을 예측에 활용

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2024	234.8063	198.7214	270.8913	179.6191	289.9935
Feb 2024	202.3298	164.0108	240.6488	143.7260	260.9337
Mar 2024	247.1007	206.6710	287.5305	185.2688	308.9327
Apr 2024	247.8481	205.4125	290.2838	182.9484	312.7479
May 2024	260.2300	215.8791	304.5809	192.4011	328.0589
Jun 2024	258.3822	212.1954	304.5690	187.7455	329.0189
Jul 2024	269.5676	221.6150	317.5201	196.2305	342.9046
Aug 2024	259.2515	209.5960	308.9069	183.3100	335.1929
Sep 2024	273.3729	222.0710	324.6748	194.9134	351.8323
Oct 2024	235.7042	182.8071	288.6013	154.8051	316.6034
Nov 2024	238.8926	184.4470	293.3382	155.6253	322.1600
Dec 2024	238.9907	183.0395	294.9420	153.4207	324.5608



2024년 6월 월간 판매량은?

2024년 4월 총선 → 6월 담배값 8000원 인상 가정

SARIMA 예측 판매량:
약 2.58억갑

DID 처치효과:
장기적 관점
통계적 유의 X

RDD 처치효과:
단기적 관점 유의미
대폭 감소 (LATE)



2024년 6월 월간 판매량은?

2024년 4월 총선 → 6월 담배값 8000원 인상 가정

SARIMA 예측 판매량:
약 2.58억갑

DID 처리효과:
통계적 유의 X

RDD 처리효과:
대폭 감소 (LATE)

2024년 6월 최종 예측 담배 판매량 :
2.58억갑 대비 대폭 감소할것

