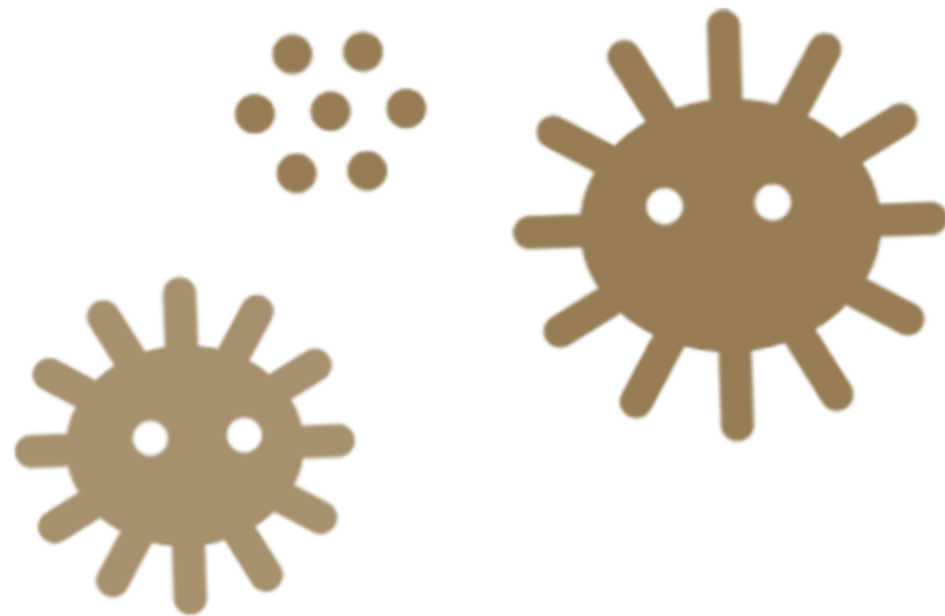


| 비모수통계학을 활용한 미세먼지 및 초미세먼지 동향 분석 |



통계학과 2019110461

최대한

통계학과 2019110477

김부겸

목 차

1
미세먼지의 정의 및 원인

2
연구의 자료 및 방법

3
미세먼지와 초미세먼지의 관계

4
코로나 전후 미세먼지 농도 변화 분석

5
미세먼지 저감조치 효과 확인

6
결론

7
이외의 내용

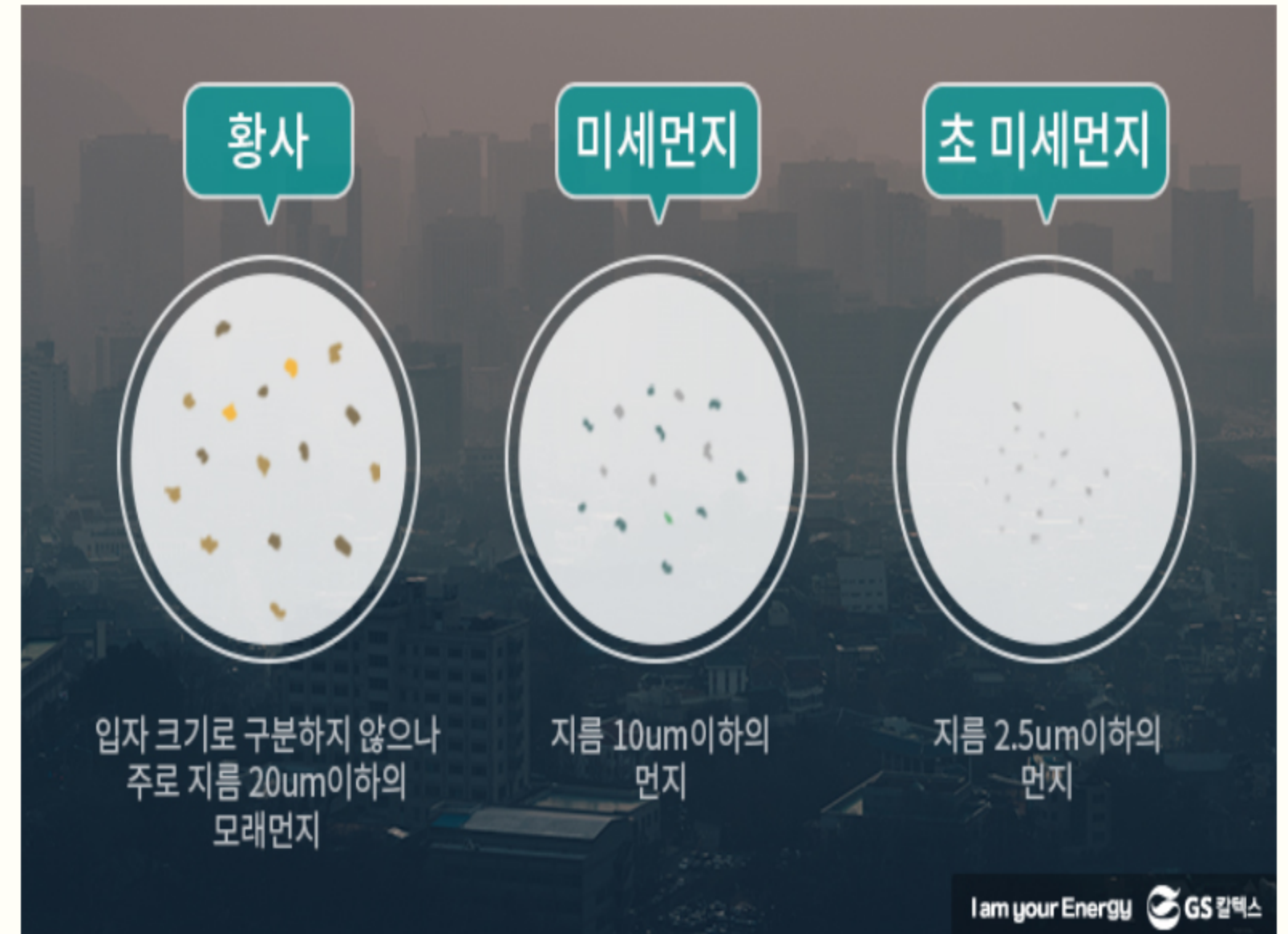
미세먼지 정의와 원인

정의

- 대기 중에 떠다니면 눈에 보이지 않을 정도로 작은 먼지
- 질산염, 암모늄 이온, 황산염 등의 이온 성분과 탄소 화합물과 금속 화합물 등으로 이루어짐.
- 먼지의 지름 크기에 따라 미세먼지와 초미세먼지로 구분.

원인

- 화학반응 또는 자연적으로 생성
- 중국에서 배출되는 오염물질
- 여러 질병과 연계되어 위협적임.



연구 자료

미세먼지 데이터

 이 누리집은 대한민국 공식 전자정부 누리집입니다.

검색어를 입력해 주세요.

로그인 회원가입 사이트맵 ENGLISH

DATA 공공데이터포털 .GO.KR

데이터찾기 국가데이터맵 데이터요청 데이터활용 정보공유 이용안내

파일데이터 상세

CSV 서울특별시_시간별 (초)미세먼지

다운로드

서울특별시 대기질 자료(초미세먼지, 미세먼지) 입니다.
2008년 1월부터 2022년 12월 31일까지의 자료로 자치구별 시간 평균 자료(서울시 평균 자료 포함)입니다.
본 측정자료는 확정 전 실시간 자료입니다. 참고용으로만 활용 가능하며, 행정 목적으로는 사용할 수 없습니다.

자료 문의 : 대기정책과 02-2133-3665

0

0

관심

	A	B	C	D
1	일시	구분	미세먼지(PM10)	초미세먼지(PM2.5)
2	2022-12-31 23:00	평균	59	46
3	2022-12-31 23:00	강남구	57	44
4	2022-12-31 23:00	강동구	68	55
5	2022-12-31 23:00	강북구	59	42
6	2022-12-31 23:00	강서구	62	40
7	2022-12-31 23:00	관악구	57	38
8	2022-12-31 23:00	광진구	54	46
9	2022-12-31 23:00	구로구	64	43
10	2022-12-31 23:00	금천구	43	35

2008년 1월부터 2022년 12월 31일까지의 자치구별 미세먼지 농도자료

분석 방향

초미세먼지와 다른 대기질 성분의 상관성

> cor(a,a1,method='spearman')	cor(초미세먼지, 이산화질소)
[1] 0.3789954	
> cor(a1,a2,method='spearman')	cor(초미세먼지, 미세먼지)
[1] 0.6073059	
> cor(a1,a3,method='spearman')	cor(초미세먼지, 오존)
[1] -0.1598174	
> cor(a1,a4,method='spearman')	cor(초미세먼지, 아황산가스)
[1] 0.369863	
> cor(a1,a5,method='spearman')	cor(초미세먼지, 일산화탄소)
[1] 0.4611872	

=> **미세먼지와 초미세먼지를 위주로 분석 진행**

분석 방법

데이터 수집

미세먼지와 초미세먼지 평균 농도 데이터를 수집



데이터 전처리

R 프로그램을 이용하여 데이터를 전처리



통계 분석

여러 비모수 통계학으로 데이터 분석



가설 검정 및 결론

가설 검정 및 결론



미세먼지와 초미세먼지의 상관관계

데이터 추출

서울특별시 대기환경정보

대기질 현황

대기오염 예보·경보

대기질 통계

정보마당

알림마당

🏠

대기질 통계

>

기간별 통계

>

월별

월별 평균

측정기간

2012

측정물질

미세먼지

선택된 지역구 : 전체

전체

강남구

강동구

강북구

강서구

관악구

광진구

구로구

금천구

동대문구

동작구

마포구

서대문구

서초구

성동구

성북구

송파구

양천구

은평구

종로구

중구

중랑구

검색

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
구분	1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월		
서울시 평균	43	60	50	47	51	52	40	28	22	27	33	42	41	
강남구	43	64	50	48	53	52	41	31	25	32	37	44	43	
강동구	39	58	48	44	50	49	38	27	20	25	32	41	40	
강북구	41	52	49	46	50	49	48	37	28	32	29	38	33	
강서구	44	66	56	52	57	56	40	27	23	29	34	44	42	
관악구	40	58	51	47	49	52	40	27	20	27	31	40	40	
광진구	41	62	52	47	54	51	37	29	20	25	33	42	41	
구로구	41	57	50	48	50	53	41	26	21	28	30	43	43	
금천구	43	58	53	48	51	56	47	33	30	32	32	34	42	
노원구	41	62	49	38	49	52	42	32	23	25	30	42	43	
도봉구	38	58	49	44	48	47	37	25	19	23	31	39	41	
동대문구	43	63	54	49	54	53	41	28	22	27	34	44	43	
동작구	43	51	46	48	51	49	41	32	27	32	37	49	48	
마포구	43	66	55	48	53	53	39	27	21	27	34	44	43	
서대문구	38	59	47	44	47	48	36	24	18	23	29	40	39	
서초구	43	65	54	50	55	55	42	28	22	28	35	45	41	
성동구	43	61	53	48	54	59	47	26	16	27	35	44	44	
성북구	38	56	47	43	48	45	37	25	19	23	31	39	38	
송파구	39	60	50	45	49	47	36	25	18	25	33	40	38	
양천구	42	59	49	47	49	55	42	30	24	33	37	42	42	
영등포구	44	69	54	50	55	57	40	30	22	34	36	47	46	

월별 미세먼지와 초미세먼지의 평균

```

> dataall$y
[1] 28 29 43 55 57 57 60 58 63 38 38 29 29 33 45 46 49 84 71 45 45 35 30 34 29 45 33 48 50 45 64 71 56 46 33 34 37 37 53
[40] 48 53 46 60 56 63 41 33 21 32 29 42 50 52 53 52 52 42 37 24 22 20 28 52 43 66 57 69 41 53 29 26 25 21 33 41 41 43 41
[79] 45 44 35 36 21 23 19 33 42 42 38 48 67 42 61 33 24 21 15 27 46 39 44 40 42 45 33 22 25 20 22 25 39

> dataall$y1
[1] 15 22 31 29 31 28 29 27 22 24 18 16 18 22 21 25 30 30 21 22 20 19 22 16 25 21 28 27 23 32 30 29 28 22 22 24 20 28 30
[40] 32 28 40 26 24 22 22 13 19 15 22 31 32 30 34 26 23 24 16 13 10 15 28 24 38 35 45 21 29 20 19 16 12 15 20 28 29 28 25
[79] 21 19 21 13 15 11 17 24 27 21 29 33 19 20 19 15 13 7 13 27 23 29 26 21 22 17 12 16 12 11 14 23 19

\

```

2013년 10월부터 2022년 12월 까지 111개의 데이터

미세먼지와 초미세먼지의 상관관계

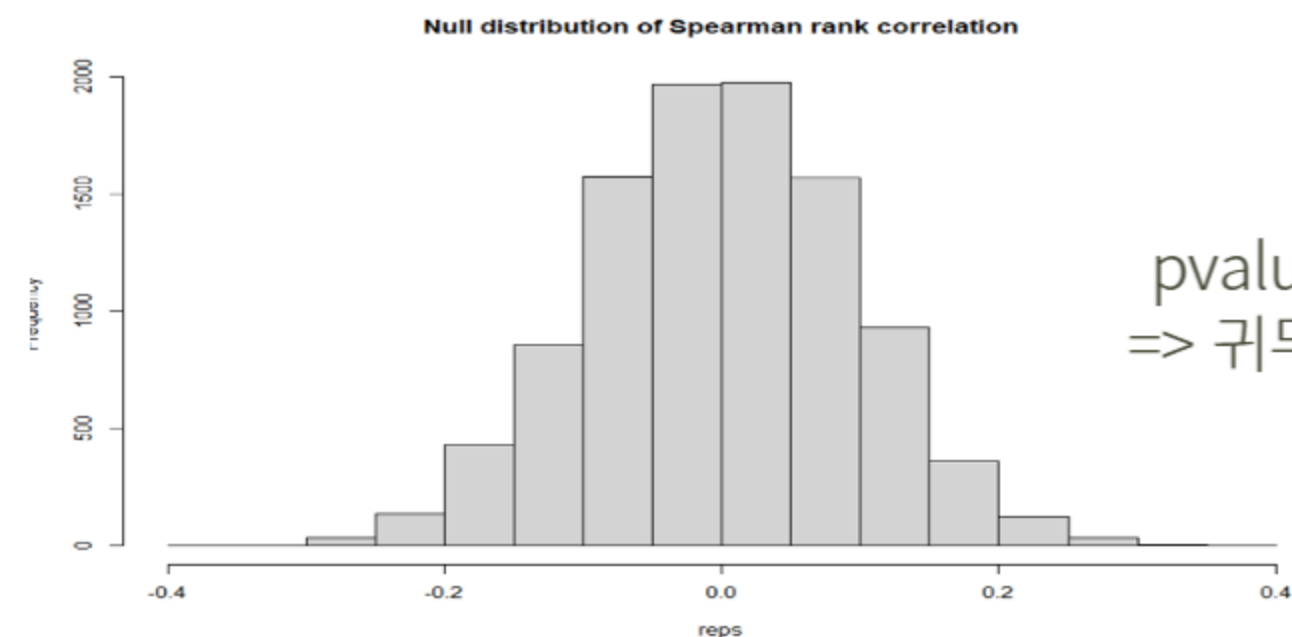
"미세먼지와 초미세먼지의 상관관계가 없다." vs "미세먼지와 초미세먼지의 상관관계가 있다."

Spearman rank correlation

```
> cor(data14$Y,data141$Y1,method="spearman")
[1] 0.8074205
> cor(data14$Y,data141$Y1,method="spearman")
[1] 0.8074205
> cor(data15$Y,data151$Y1,method="spearman")
[1] 0.9113475
> cor(data16$Y,data161$Y1,method="spearman")
[1] 0.8697237
> cor(data17$Y,data171$Y1,method="spearman")
[1] 0.8450914
> cor(data18$Y,data181$Y1,method="spearman")
[1] 0.9270355
> cor(data19$Y,data191$Y1,method="spearman")
[1] 0.9418121
> cor(data20$Y,data201$Y1,method="spearman")
[1] 0.8175439
> cor(data21$Y,data211$Y1,method="spearman")
[1] 0.8736896
> cor(data22$Y,data221$Y1,method="spearman")
[1] 0.9050981
```

Permutation test

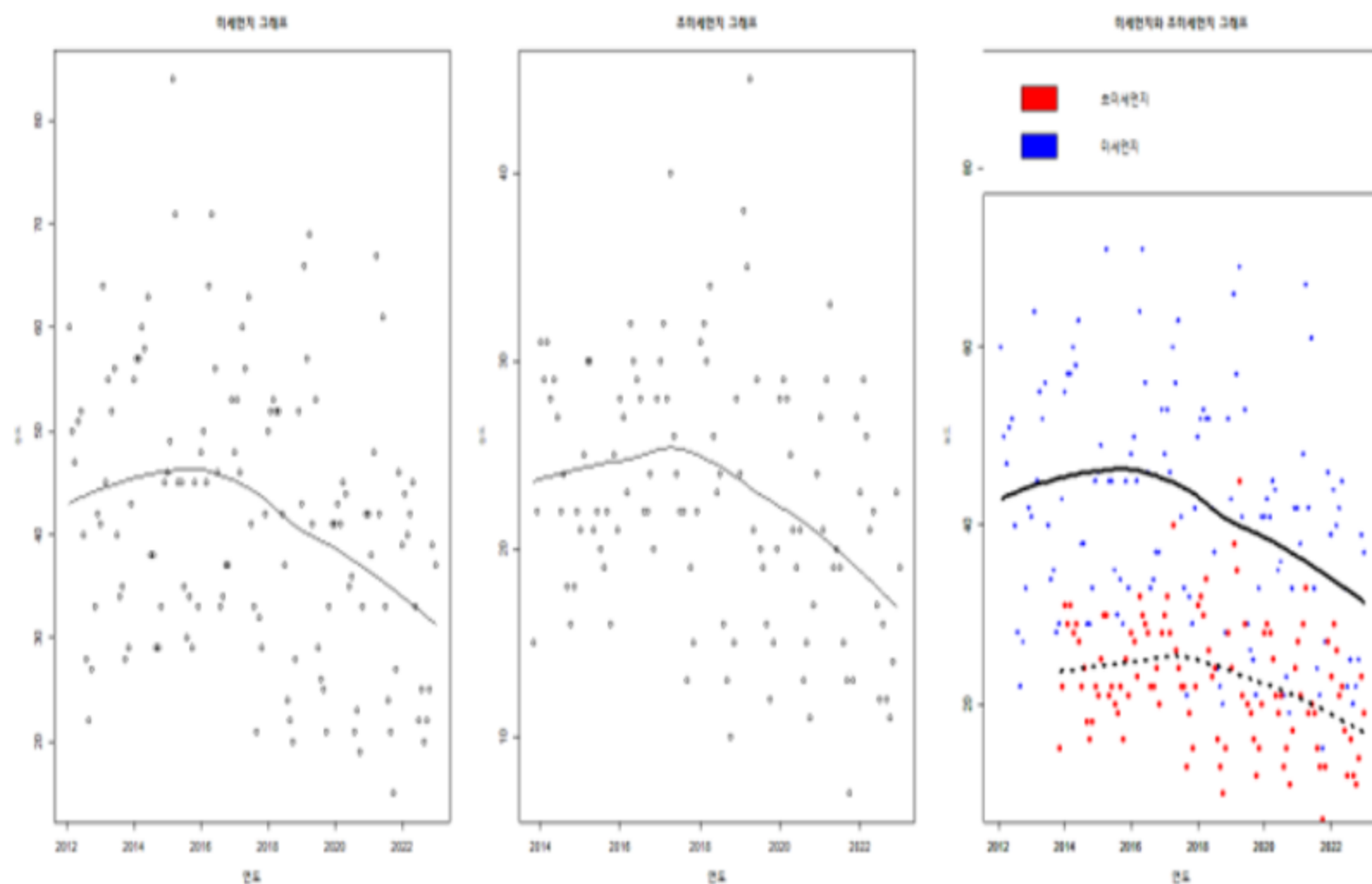
```
> par(mfrow=c(1,1))
> R <- 9999
> reps <- numeric(R)
> # calculate the observed spearman rank correlation
> spearmancorr <- cor(dataall$Y,dataall$Y1, method="spearman")
> spearmancorr
[1] 0.6241997
> for (i in 1:R) {
+   x <- dataall$Y
+   y <- sample(dataall$Y1, size=dim(dataall)[1], replace=FALSE)
+   reps[i] <- cor(y,x, method="spearman")
+ }
> pvalue <- mean(c(spearmancorr, reps) >= spearmancorr)
> pvalue
[1] 1e-04
> hist(reps, main="Null distribution of Spearman rank correlation")
> abline(v=spearmancorr, lty=2 ,col="red")
```



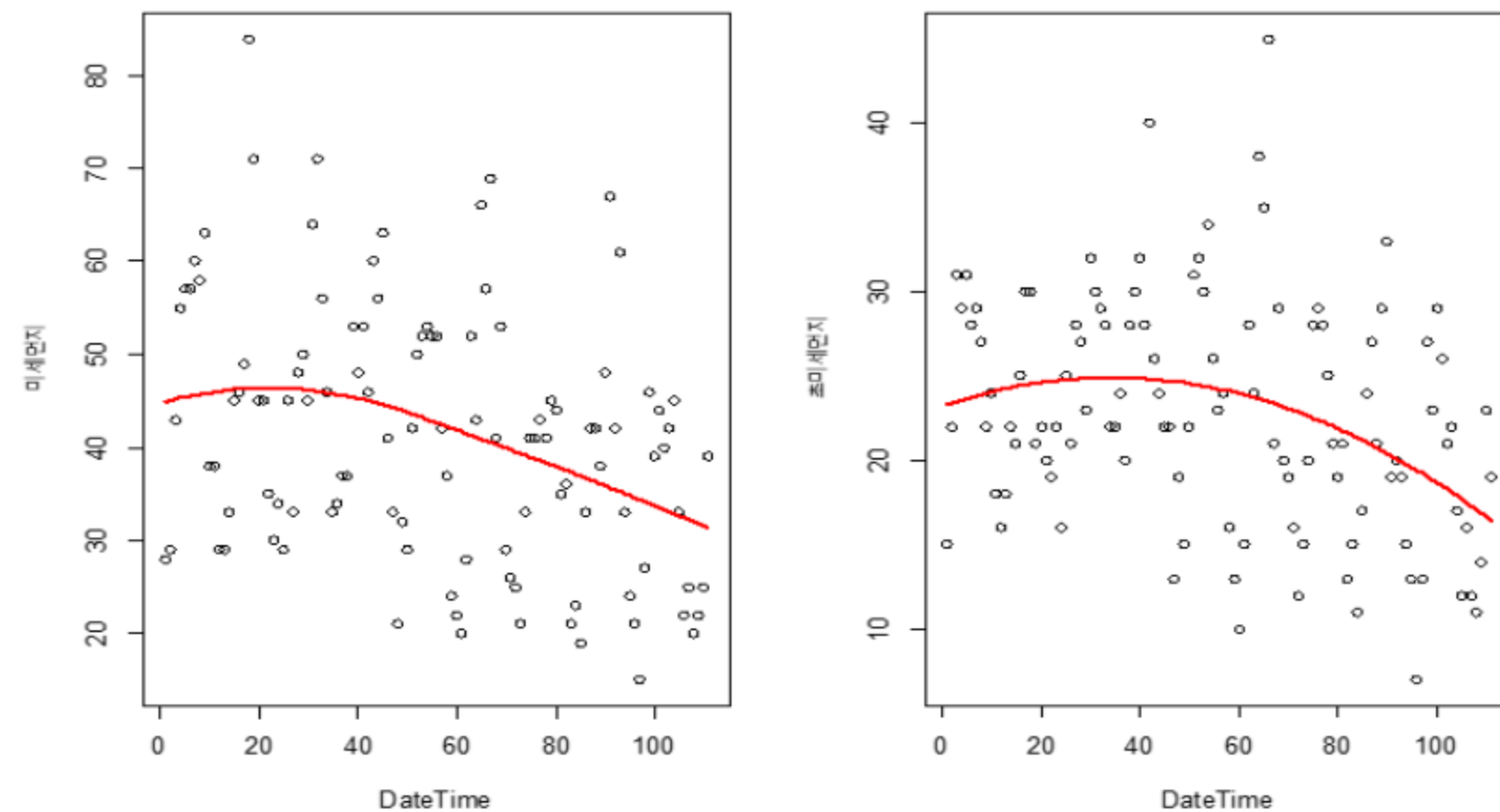
pvalue=0.0001
=> 귀무가설 기각

미세먼지와 초미세먼지의 상관관계

LOESS nonparametric regression



Regression Splines



비슷한 추세를 보이고 있음.

코로나 전후 미세먼지 농도 변화 분석

귀무 가설

"코로나 발생 전과 발생후의 평균 미세먼지 농도는 같다."

VS

대립 가설

"코로나 발생 전이 발생 후보다 평균 미세먼지 농도가 높다."

```
> before
```

```
[1] 41.04376 44.58172 45.93256 45.59775 47.86019 43.85726 39.78469 42.15119
```

```
> after
```

```
[1] 35.29584 38.24855 32.86100
```

코로나 발생년도인 2020년을 기준
2012년~2019년 => before
2020년~2022년 => after

코로나 전후 미세먼지 농도 변화 분석

Unpaired Data

Wilcoxon rank sum test

```
> wilcox.exact(before, after, paired=FALSE, alternative="greater")
```

Exact Wilcoxon rank sum test

data: before and after

W = 24, p-value = 0.006061

alternative hypothesis: true mu is greater than 0

pvalue=0.006061

=> 귀무가설 기각

Permutation test

```
> R <- 9999
> ranksall <- rank(c(before,after))
> k <- 1:length(ranksall)
> reps <- numeric(R)
> ts <- sum(ranksall[1:length(before)])
> ts
[1] 60
> for (i in 1:R) {
+   m <- sample(k, size=length(before), replace=FALSE)
+   a1 <- ranksall[m]
+   reps[i] <- sum(a1)
+ }
> pvalue <- mean(c(ts, reps) >= ts)
> pvalue
[1] 0.0062
```

pvalue=0.0062

=> 귀무가설 기각

"코로나 발생 전이 발생 후보다 평균 미세먼지 농도가 높다."

코로나 전후 미세먼지 농도 변화 분석

귀무 가설

"코로나 발생 전과 발생후의 평균 미세먼지 농도의 중앙값이 같다."

VS

대립 가설

"코로나 발생 전이 발생 후보다 평균 미세먼지 농도의 중앙값보다 크다"

```
> apply(df1219,1,mean)
강남구  강동구  강북구  강서구  관악구  광진구  구로구  금천구
44.05621 44.90970 41.89617 45.81310 46.40042 44.49963 45.41598 42.66234
노원구  도봉구  동대문구  동작구  마포구  서대문구  서초구  성동구
42.22216 42.07749 43.33755 42.44855 45.26285 43.27884 46.03540 45.92097
성북구  송파구  양천구  영등포구  용산구  은평구  종로구  중구
43.38160 42.62514 44.97978 47.72614 40.19078 43.50653 42.54607 41.67033
중랑구
43.41480
> apply(df2022,1,mean)
강남구  강동구  강북구  강서구  관악구  광진구  구로구  금천구
33.67992 36.69970 37.32195 38.37459 34.99903 34.06177 37.31113 34.95798
노원구  도봉구  동대문구  동작구  마포구  서대문구  서초구  성동구
35.18153 32.83383 34.81602 38.64212 32.82103 32.83850 36.89669 35.13475
성북구  송파구  양천구  영등포구  용산구  은평구  종로구  중구
36.04789 35.27069 37.86697 34.54936 35.48213 34.41258 35.54077 35.81494
중랑구
35.15577
```

코로나 발생년도인 2020년을 기준으로 자치구별 평균
2012년~2019년 자치구별 평균=> before
2020년~2022년 자치구별 평균=> after

코로나 전후 미세먼지 농도 변화 분석

Paired Data

Sign test

```
> before<-apply(df1219,1,mean)
> after<-apply(df2022,1,mean)
>
> d<-before-after
> T<-length(d[d>0])
> binom.test(T, length(d[d != 0]), alternative="greater")
```

Exact binomial test

```
data: T and length(d[d != 0])
number of successes = 25, number of trials = 25, p-value = 2.98e-08
alternative hypothesis: true probability of success is greater than 0.5
95 percent confidence interval:
 0.8870719 1.0000000
sample estimates:
probability of success
1
```

pvalue=2.98e-08
=> 귀무가설 기각

Signed-Ranks Test

```
> library(exactRankTests)
> wilcox.exact(x, y, paired=TRUE, alternative="greater")
```

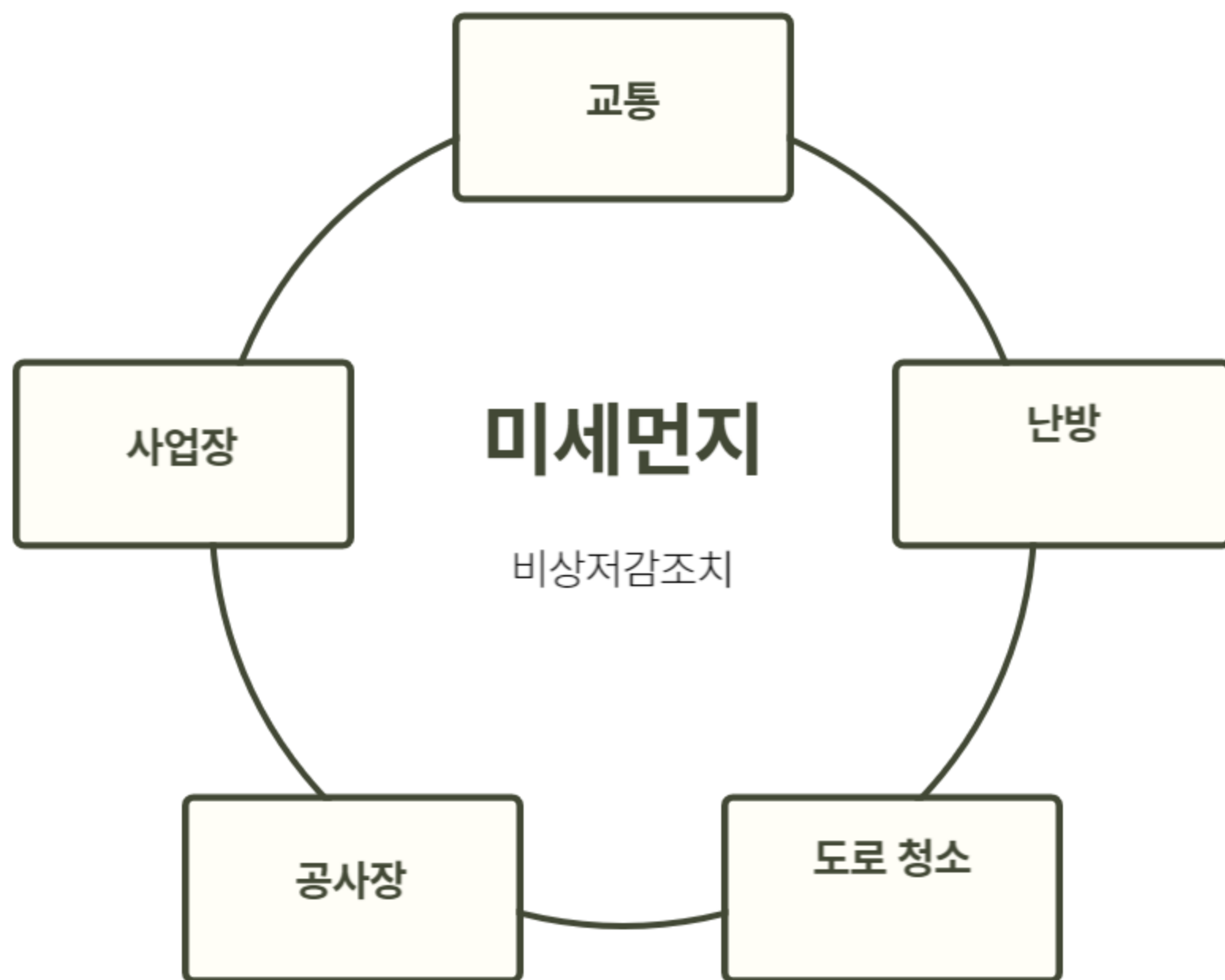
Exact wilcoxon signed rank test

```
data: x and y
V = 325, p-value = 2.98e-08
alternative hypothesis: true mu is greater than 0
```

pvalue=2.98e-08
=> 귀무가설 기각

"코로나 발생 전이 발생 후보다 평균 미세먼지 농도의 중앙값보다 크다"

미세먼지 비상저감조치란?



서울 지역 내 고농도 미세먼지(PM-2.5)가 일정기간 지속될 경우, 시민건강을 위해 미세먼지를 단기간에 줄이고자 자동차, 공장, 공사장의 대기오염 물질을 줄이는 조치

비상저감조치 시행일시 : 발령일 다음날 06시~21시

"미세먼지 발생 원인 중 국내 내부 발생요인을 최대한 줄이자!"

미세먼지 저감조치 효과 확인

33	비상저감조치	관심	2021-03-14 17:00	2021-03-15 06:00 ~ 2021-03-15 21:00
32	예비저감조치	-	2021-03-13 17:00	2021-03-14 06:00 ~ 2021-03-14 21:00
31	비상저감조치	관심	2021-03-11 17:00	2021-03-12 06:00 ~ 2021-03-12 21:00
30	비상저감조치	관심	2021-03-10 17:00	2021-03-11 06:00 ~ 2021-03-11 21:00

비상저감조치 시행일시 : 발령일 다음날 06시~21시

ex) 2022-02-11 17시 비상저감조치 발령 => 2022-02-12 06-21시 비상저감조치 시행
2022-02-11 21시와 2022-02-12 21시 초미세먼지 데이터 비교 분석

귀무 가설

비상저감조치 전/후 시간대의 농도의 중앙값이 같다.

대립 가설

비상저감조치 전이 후보다 시간대의 농도의 중앙값이 높다.

미세먼지 저감조치 효과 확인

Paired Data

Sign Test

```
> binom.test(T, length(d[d != 0]), alternative="greater")

Exact binomial test

data:  T and length(d[d != 0])
number of successes = 25, number of trials = 25, p-value = 2.98e-08
alternative hypothesis: true probability of success is greater than 0.5
95 percent confidence interval:
 0.8870719 1.0000000
sample estimates:
probability of success
                  1
```

pvalue=2.98e-08
=> 귀무가설 기각

Signed-Ranks Test

```
> wilcox.exact(a, b, paired=TRUE, alternative="greater")

Exact Wilcoxon signed rank test

data:  a and b
V = 325, p-value = 2.98e-08
alternative hypothesis: true mu is greater than 0
```

pvalue=2.98e-08
=> 귀무가설 기각

"비상저감조치 전이 후보다 시간대의 농도의 중앙값이 높다".

결론

1.미세먼지와 초미세먼지는 유의미한 상관관계가 있다.

Spearman correlation, Permutation ,LOESS ,Spline

2.코로나 이후 미세먼지 농도가 감소하였다.

Wilcoxon rank sum test, Permutation, Sign test, Signed-Ranks Test

3.미세먼지 비상저감조치가 효과가 있다.

sign test , wilcoxon sign rank test



감사합니다!