

미세먼지 농도에 따른 서울시 교통량 분석

프로젝트형 빅데이터 분석 서비스 개발
R 미니 프로젝트



목 차

01

프로젝트 개요

주제 선정 이유

팀 구성 및 역할

02

데이터 분석

데이터 수집

데이터 분석

03

느낀 점

소감

CHAPTER. 1

프로젝트 개요

주제 선정 이유

한반도를 뒤덮은 미세먼지가 연일 기승을 부리며 시민들의 건강을 위협하고 있다.

실제 미세먼지 농도가 일정 기준 이상으로 높을 때 시행되는 비상저감조치

전후의 교통량 변화를 분석해보고, 이것이 대기오염 물질의 배출을

낮추는 데 얼마나 유의미한 효과를 거두는지 알아보고자 해당 주제를 선정했다.





중앙일보 | 4일 전 | 네이버뉴스

시민이 본 서울시 "최고는 대중교통, 최악은 미세먼지"



아주경제 | 2021.03.15.

"오늘도 미세먼지 비상저감조치 합니다"

전기연 기자 kiyeoun01@ajunews.com "오늘도 미세먼지 비상저감조치 합니다" 14



한겨레 PICK | 2021.03.14. | 네이버뉴스

오늘 외출 조심하세요...수도권 전역 **미세먼지 나쁨**

미세먼지 예비저감조치는 이들 후 비상저감조치가 시행될 가능성이 높을 경우 사

비상저감조치란?

고농도 미세먼지(PM_{2.5})가 일정기간 지속시 국민건강을 보호하기 위해
단기간에 대기질을 개선할 수 있는 조치를 뜻합니다.



24

아이뉴스24 | 20시간 전 | 네이버뉴스

[지금은 기후위기] **미세먼지** 얼마나 고통스럽고 지겨웠으면...

그동안 국민이 제안한 미세먼지 저감 대책을 보면 '얼마나 미세먼지로 국민이 고통



뉴스1 | 52분 전 | 네이버뉴스

미세먼지로 뿌연 서울

수도권을 중심으로 전국 곳곳에 미세먼지 농도가 '나쁨' 수준을 보인 23일



공 하 영

데이터 수집 및 전처리

데이터 분석

시각화



김 대 찬

데이터 수집 및 전처리

데이터 분석

시각화

예상 가설

“미세먼지 농도가 높아지면 사람들의 외출량과
서울 시내 교통량은 줄어든 것이다.”

CHAPTER. 2

데이터 분석

데이터 수집



미세먼지 키워드
동적 웹크롤링



서울시 대기환경정보
동적 웹크롤링



지하철 이용객 정보
서울 열린데이터 광장



서울시 교통량 정보
서울교통정보센터

미세먼지 데이터 수집

측정기간

2021

1월

측정물질

초미세먼지

전체	강남구	강동구	강북구	강서구	관악구	광진구	구로구	금천구	노원구	도봉구
동대문구	동작구	마포구	서대문구	서초구	성동구	성북구	송파구	양천구	영등포구	용산구
은평구	종로구	중구	중랑구							

검색

! 이 자료는 최종확정자료가 아닙니다.
최종확정자료는 환경부 대기환경정보 홈페이지(<http://www.airkorea.or.kr>) 통계정보에서 확인

측정기간 : 2021년 01월 , 초미세먼지 PM-2.5 ($\mu\text{g}/\text{m}^3$)

엑셀다운로드

초미세먼지 PM-2.5 ($\mu\text{g}/\text{m}^3$)

구분	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
서울시 평균	21	16	11	13	21	16	13	11	10	15	22	31	47	45	29	32	35	13	18	14	24

2019년 초미세먼지 / 미세먼지 서울시 평균
2000년~2020년 미세먼지 서울시 평균

미세먼지 관련 검색어 수집



1주일간 네이버 기사 제목 네이버 실시간 검색 결과

회귀분석 (lm 결과)

```
> result.lm <- lm(초미세먼지~미세먼지+탑승객+교통량, data=result)
> summary(result.lm)
```

Call:
lm(formula = 초미세먼지 ~ 미세먼지 + 탑승객 + 교통량, data = result)

Residuals:

	Min	1Q	Median	3Q	Max
	-41.210	-1.791	0.772	3.335	24.219

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.728e+01	7.215e+00	2.396	0.0171 *
미세먼지	6.383e-01	1.372e-02	46.541	<2e-16 ***
탑승객	5.531e-07	3.947e-07	1.401	0.1619
교통량	-2.564e-06	1.096e-06	-2.339	0.0199 *

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

Residual standard error: 6.738 on 361 degrees of freedom
Multiple R-squared: 0.858, Adjusted R-squared: 0.8568
F-statistic: 727.2 on 3 and 361 DF, p-value: < 2.2e-16

연도별 미세먼지 평균 농도



2000년대 초반
미세먼지

35~45

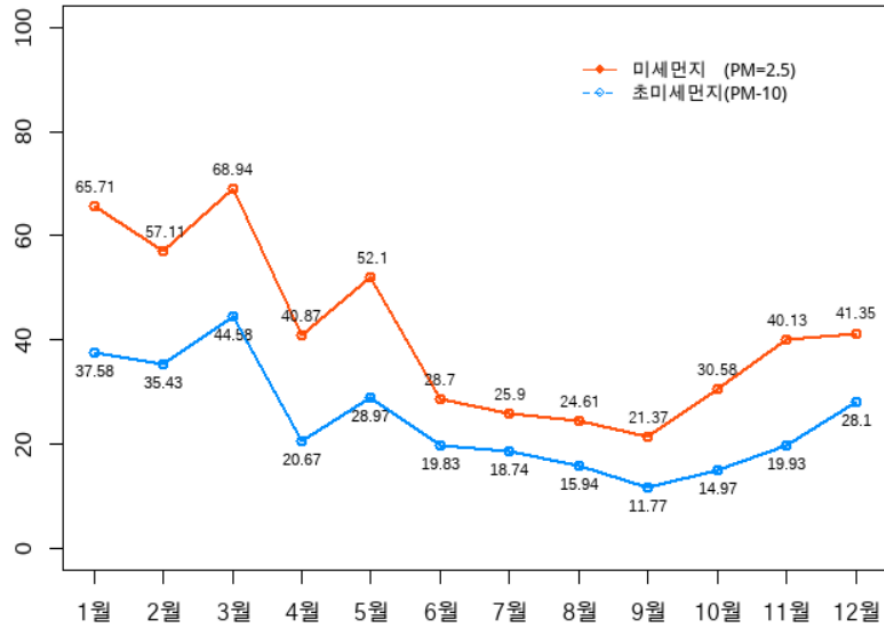


2010년대 후반
미세먼지

65~75

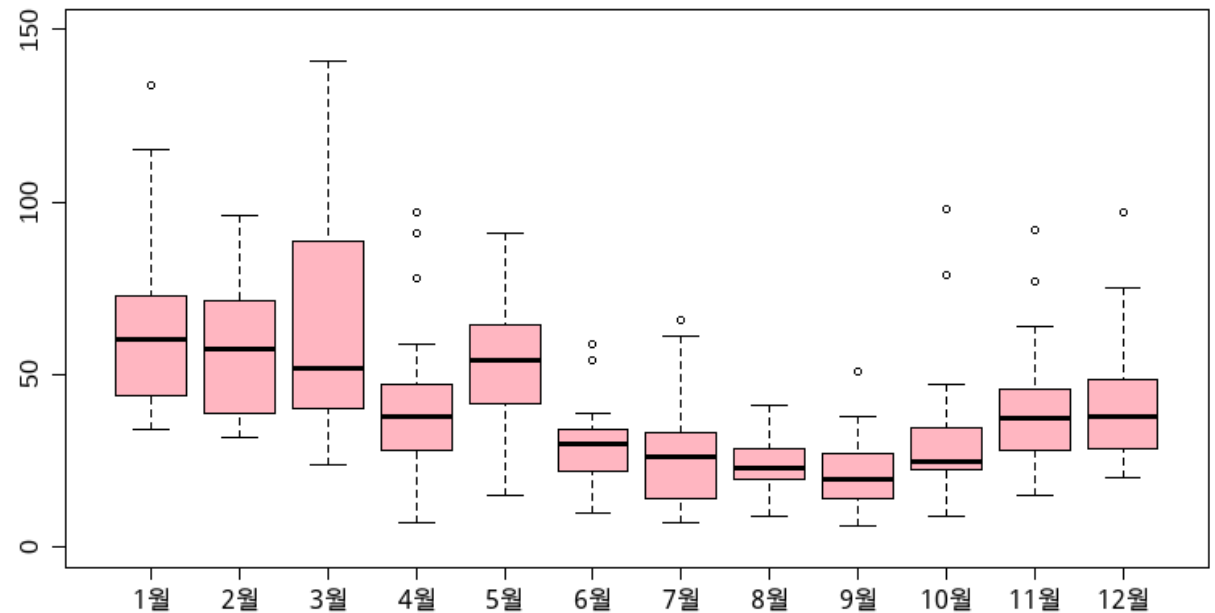
월별 미세먼지 평균 농도

2019년 초/미세먼지 월별 농도 변화



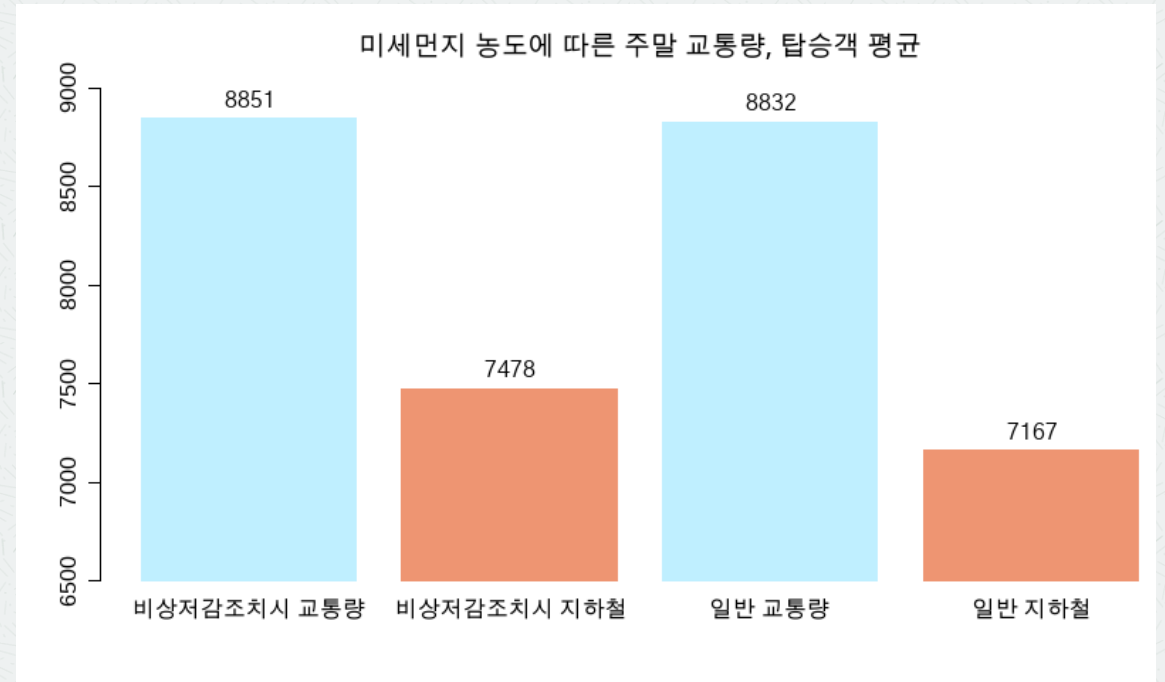
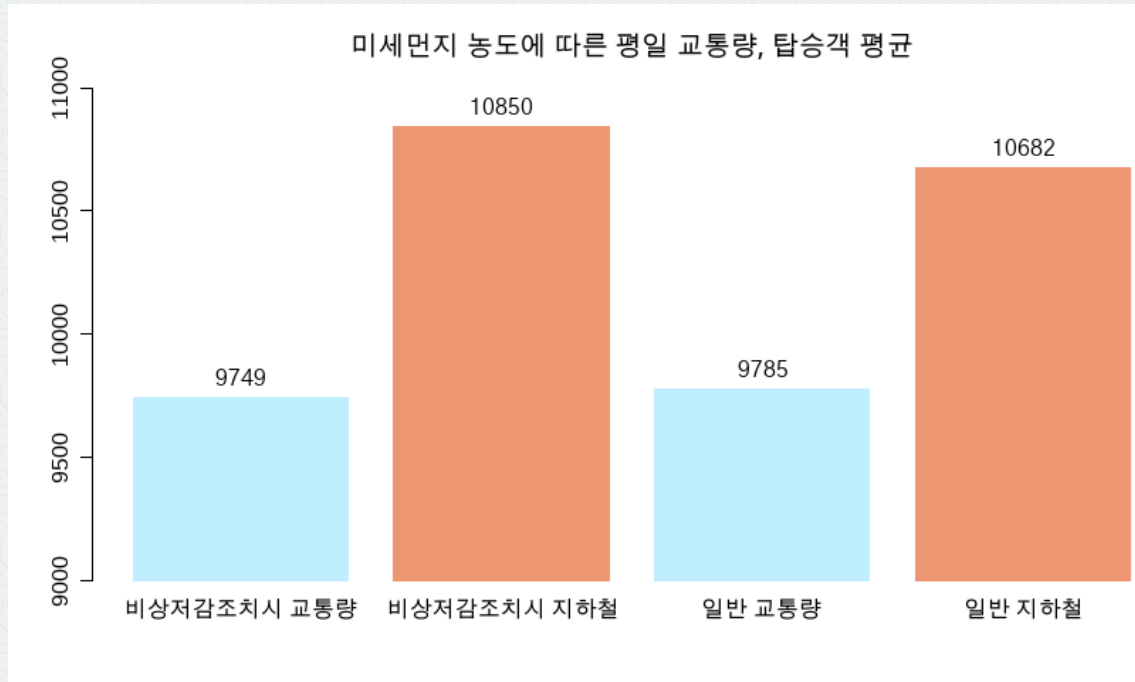
봄철 편서풍의 영향을 많이 받는 시기에
미세먼지 농도가 높음

2019년 미세먼지 월별 농도 변화



미세먼지는 다양한 요인에 발생하기
때문에 이상치가 많음

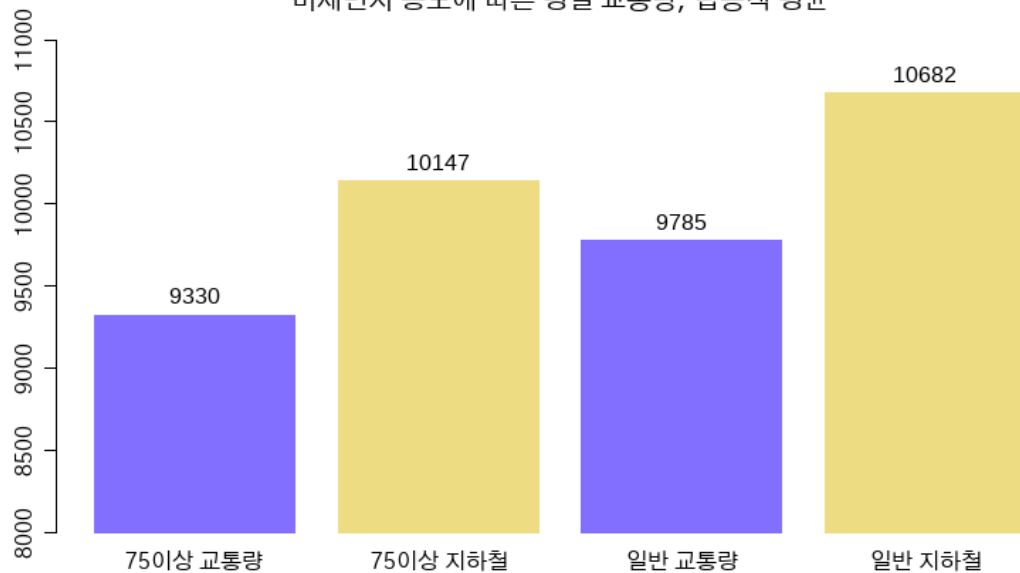
미세먼지농도 50 μ g



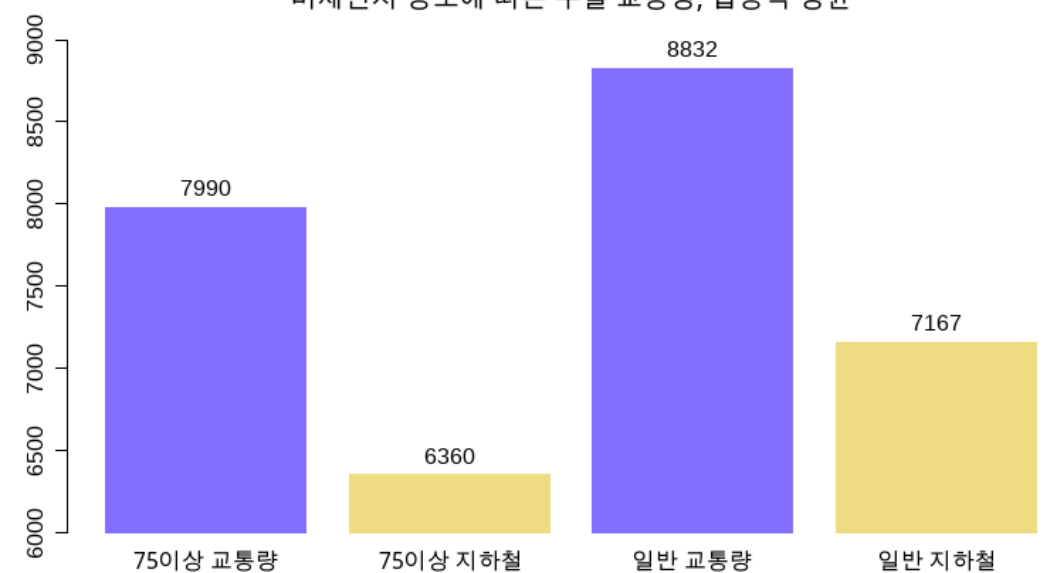
평일과 주말에 따른 차이는 보이지만
미세먼지와는 관계 없음

미세먼지농도 75 μ g

미세먼지 농도에 따른 평일 교통량, 탑승객 평균

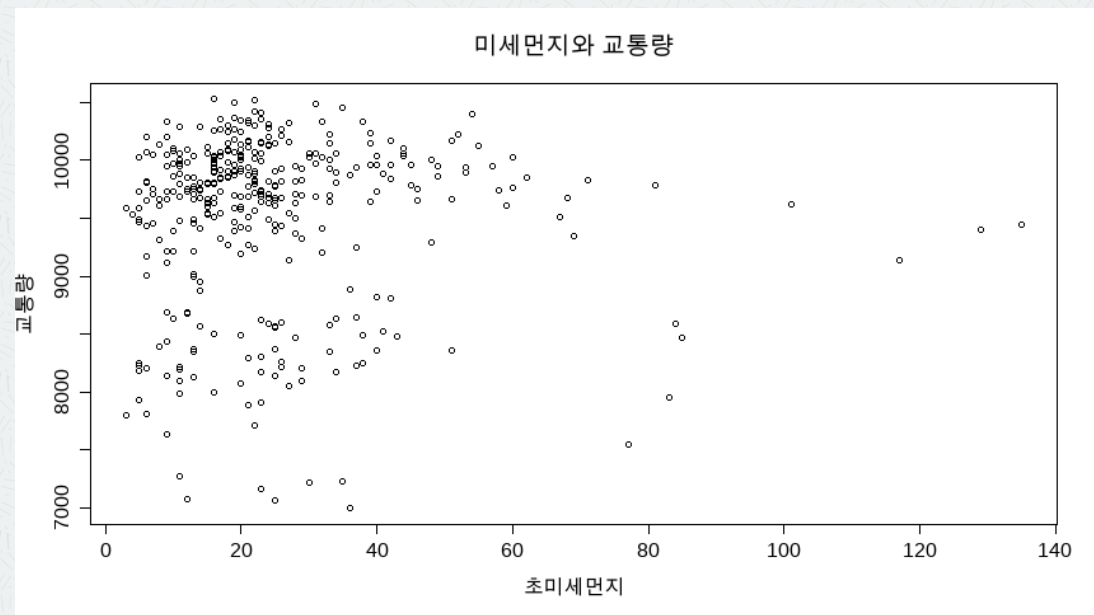
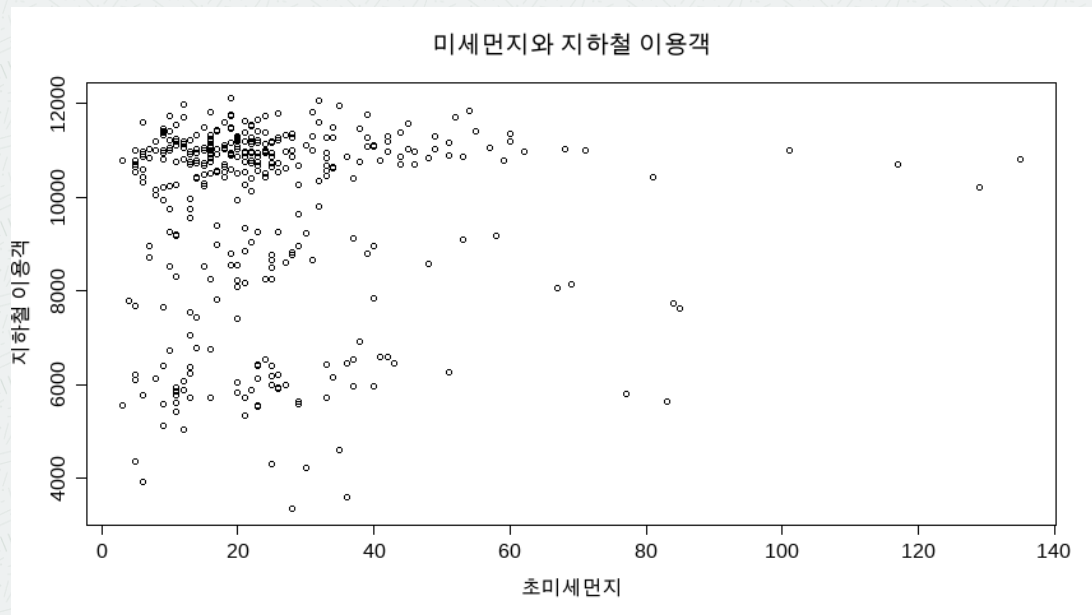


미세먼지 농도에 따른 주말 교통량, 탑승객 평균



미세먼지가 아주 심할 경우에는 감소

결론



산포도에서 볼 수 있듯이 미세먼지와 외출은 큰 상관이 없다.

CHAPTER. 3

느낀 점

공하영

기존 데이터를 목적에 맞게 수집, 처리하고 또 분석한다는 것이 얼마나 어려운 일인지 체감할 수 있었습니다. 미니 프로젝트를 경험하며, 우리가 예상했던 가설과 달라 유의미한 결과를 도출하지 못했다는 점이 아쉽지만, 이번 프로젝트를 바탕으로 빅데이터 분석을 수월하게 진행 수 있는 초석으로 작용할 것 같습니다.

김대찬

내가 필요로 하는 데이터를 수집하고 그것을 전처리하는 과정까지 생각보다 어려움이 많아 고생을 꽤 했던 것 같습니다. 하지만, 배운 것들을 활용해 볼 수 있는 좋은 기회였고 많은 것을 배웠다고 생각합니다. 또한, 그동안 생각만 하던 것을 실제로 검증해 볼 수 있던 좋은 기회였습니다.

THANK YOU