

광진구

1. 연도별 평균 데이터를 이용해 전체적인 추이를 보자!

분석에 사용한 데이터: 서울시 연도별 평균 대기오염도 정보.csv

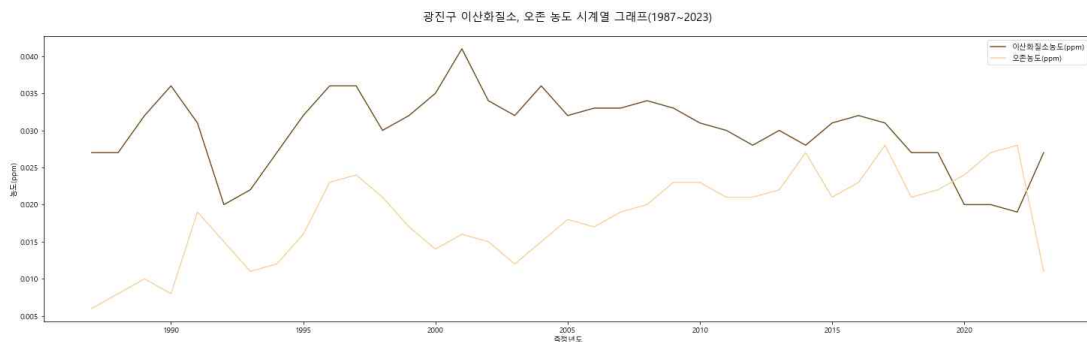
결측치

- 미세먼지 13개
- 초미세먼지 25개

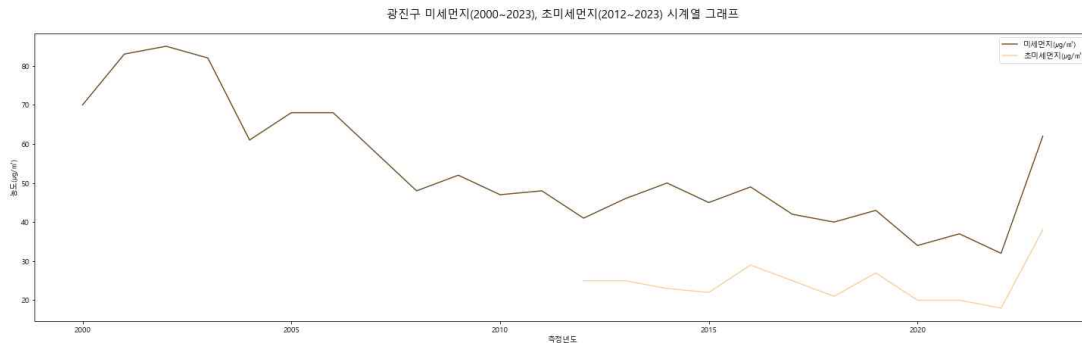
관측 기간

- 이산화질소, 오존: 1987~2023년
- 미세먼지: 1995년~2023년
- 초미세먼지: 2012~2023년

연도별 데이터의 경우 미세먼지, 초미세먼지에만 결측치가 존재한다. 이는 미세먼지, 초미세먼지를 제외한 다른 대기오염물질의 경우 측정이 1987년부터 시작되었는데, 미세먼지는 1995년부터, 초미세먼지는 2012년부터 측정 데이터가 존재하여 발생한 결측치이다. 따라서 보간할 필요가 없다고 판단하여 시각화할 때 시간축 범위만 조정하고 그냥 두었다. 다만 미세먼지의 경우 첫 측정이 이루어진 것은 1995년이나, 모든 측정소에서 측정이 시작된 것이 아니라서 데이터가 다소 부족해 2000년 데이터부터 사용하였다.



광진구의 1987~2023년 범위 이산화질소, 오존 농도의 경우 해마다 불규칙적으로 변동하는 양상을 확인할 수 있었으며, 뚜렷한 경향성을 발견하지는 못했다.



광진구의 미세먼지(2000~2023년 범위)의 경우 2000년대 초반부터 꾸준히 감소하는 추세를 보이다가 최근 가파른 증가세를 보이고 있다. 초미세먼지는 2012년~2023년 범위의 데이터가 존재하는데, 그 경향성은 미세먼지와 거의 일치한다. 미세먼지 농도가 2000년대 초반부터 꾸준히 감소하는 이유는 미세먼지에 대한 국민들의 관심도 상승과 함께 미세먼지 저감조치, 미세먼지 배출원 집중관리 등 관련 정책이 증가했기 때문으로 보인다. 이러한 전체적인 감소세와는 반대로, 최근 보이는 미세먼지와 초미세먼지의 가파른 증가세의 원인은 자양 1구역 재개발 공사에 의한 것으로 추측된다.

실제로 서울시는 광진구의 미세먼지가 가파른 증가세를 보이던 2021년, 자치구 수요조사(2021년 3월), 전문가 현장평가(2021년 4월), 선정위원회 심의(2021년 5월) 및 선정지역 주민 의견수렴, 환경부 협의를 거쳐 2021년 7월, 광진구를 ‘미세먼지 집중관리구역’으로 신규 지정한 바 있다. 각 단계에서 어떠한 과정을 거쳤는지까지는 정확히 파악할 수 없었지만, 서울시가 광진구의 미세먼지 농도가 급격히 증가하는 시기에 해당 구를 미세먼지 집중관리 대상으로 지정했다는 것은 광진구의 미세먼지 관리 필요성이 검증되었다는 방증일 것이다.¹⁾

광진구는 미세먼지 집중관리 구역 지정에 따라 지원받은 3억 원의 예산을 자양 1구역 재개발 공사장과 동서울터미널 일대를 대상으로 살수차와 분진흡입차를 확대 운영하고, 공사장을 대상으로 비산먼지 제거 장치를 대여해주는 사업을 시행할 예정이라고 밝혔다. 이는 광진구에서도 미세먼지 농도를 급격하게 증가시킨 주원인이 자양 1구역 재개발 공사장임을 인식하고 있음을 시사하는 조치로 보인다.

1) <https://news.seoul.go.kr/env/archives/513823>

2. 월별 평균 데이터를 통해 월별 경향성을 살펴보자!

분석에 사용한 데이터: 서울시 월별 평균 대기오염도 정보.csv

결측치

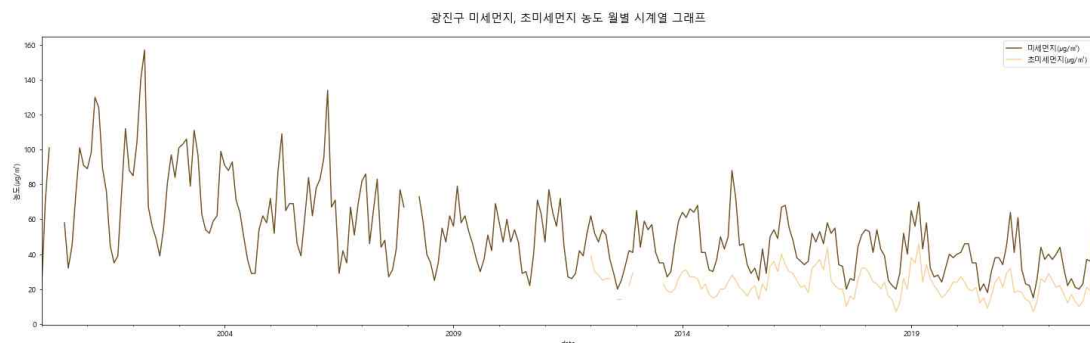
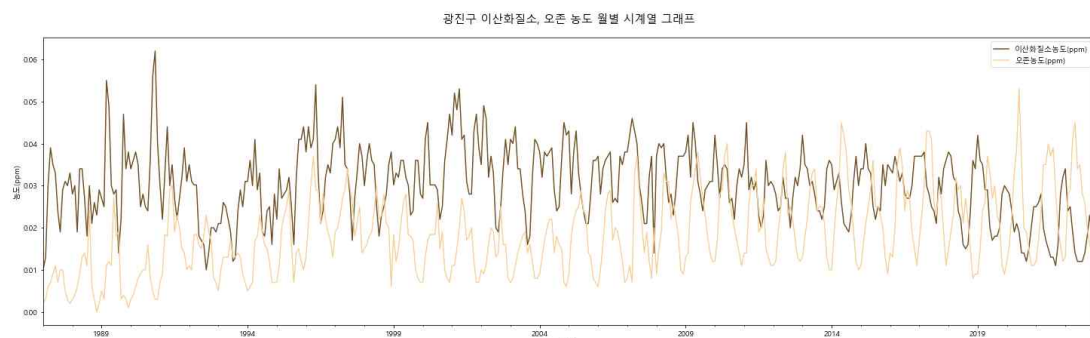
- 이산화질소 8개
- 오존 8개
- 미세먼지 162개
- 초미세먼지 308개

관측 기간

- 이산화질소, 오존: 1987.01~2023.01
- 미세먼지: 1995.01~2023.01
- 초미세먼지: 2012.01~2023.01

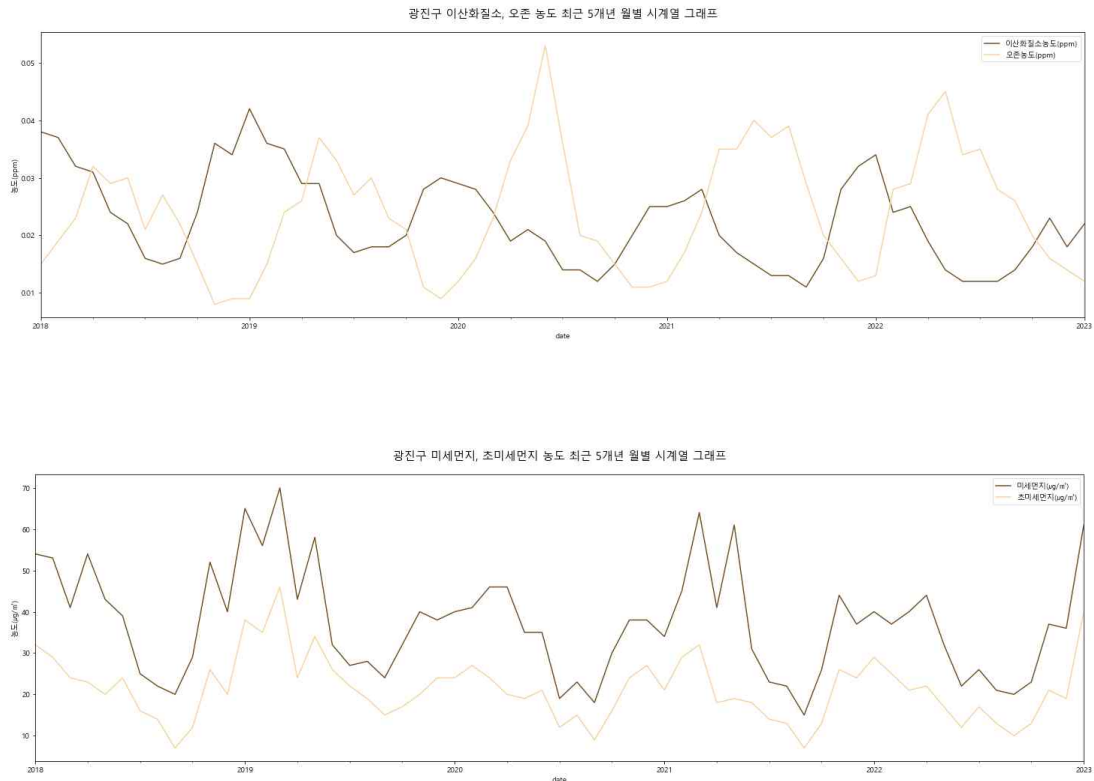
미세먼지, 초미세먼지 결측치가 많은 것은 측정 시작 시점의 차이 때문이라서 우선은 이산화질소, 오존에 대해서만 각각의 평균값을 이용해 보간하였다.

1) 전체 기간 월별 시계열 그래프



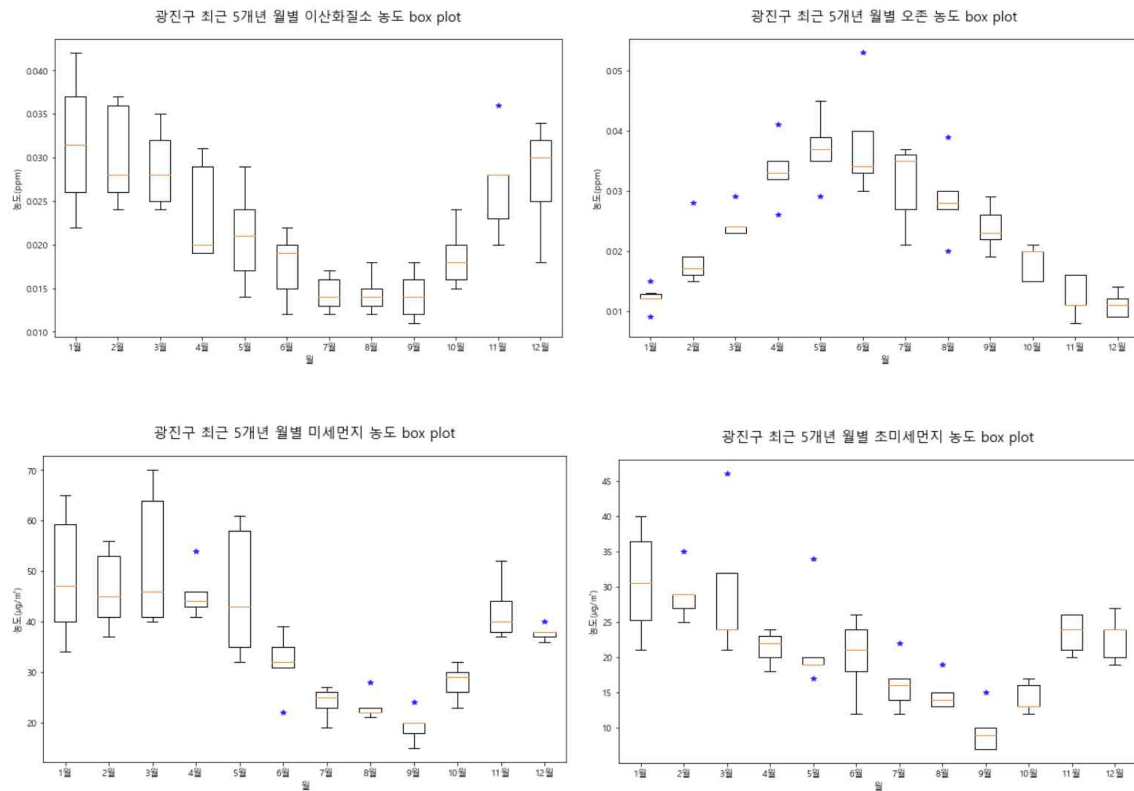
전체 기간 월별 시계열 그래프를 통해 이산화질소, 오존의 경우 서로 반대되는 경향성(이산화질소가 증가하면 오존이 감소, vice versa)을 가지며, 미세먼지, 초미세먼지 농도의 경우 연도별 평균 데이터 시각화에서 보인 전체적인 감소세를 다시 한번 확인할 수 있었다.

2) 2018~2022 최근 5개년 월별 시계열 그래프

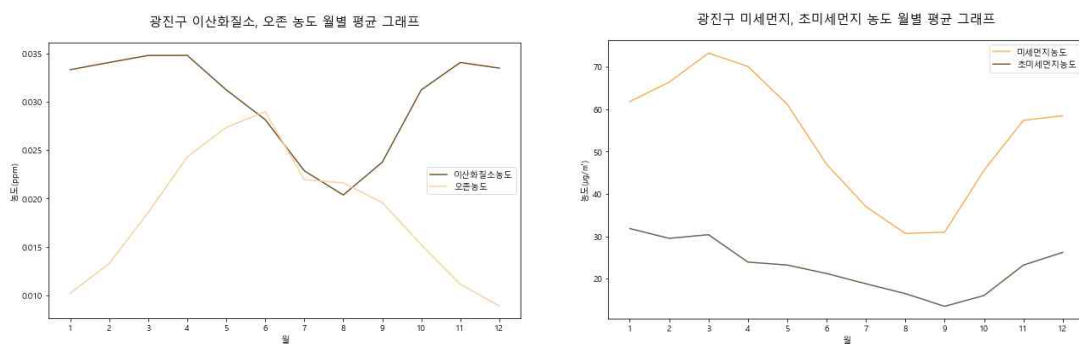


월별 변화를 보다 확실하게 보기 위하여 2018년부터 2022년까지(2023년의 경우 1월 자료밖에 없고, 아직 1월이 다 지나지 않아 그마저도 불완전한 평균 데이터이기 때문에 연도 범위에서 제외) 최근 5개년을 범위로 잡아 시각화를 다시 진행해 보았다. 그 결과 이산화질소, 오존의 서로 반대되는 경향성이 더욱 뚜렷하게 나타나며, 특히 이산화질소는 겨울에 증가, 여름에 감소, 오존은 겨울에 감소, 여름에 증가하는 계절성을 확인할 수 있었다. 미세먼지 및 초미세먼지의 경우에도 봄철에 증가하고 여름철에 감소하는 계절성을 확인할 수 있었다.

3) 최근 5개년 월별, 대기오염물질별 box plot



4) 월별 평균 그래프



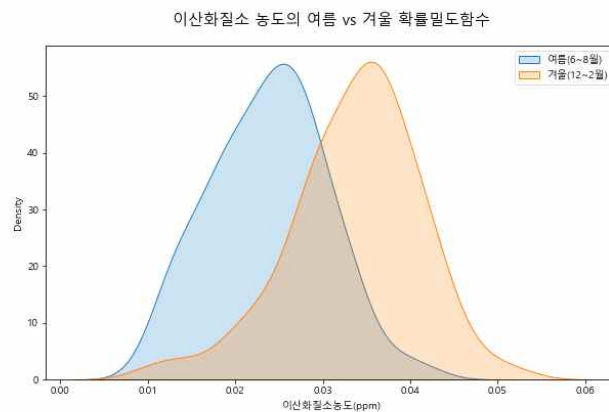
1월은 1월끼리, 2월은 2월끼리 평균을 내어 월평균 그래프를 그려본 결과 이산화질소와 오존의 계절성을 더 뚜렷하게 확인할 수 있었으며, 미세먼지와 초미세먼지의 경우에도 마찬가지로 봄철에 증가하고 여름에 감소하는 경향성을 확실하게 볼 수 있었다.

5) 이산화질소 계절성 검증을 위한 Two sample t-test

앞서 2018~2022 최근 5개년 월별 시계열 그래프 및 월별 평균 그래프를 통해 이산화질소는 겨울에 증가, 여름에 감소하는 계절성이 있는 것 같다는 가설을 이끌어낼 수 있었으므로, 이를 통계적으로 검증하고자 한다. 이에 전체 데이터(모집단)에서 6~8월 범위의 여름 표본, 12~2월 범위의 겨울 표본을 추출해 Two sample t-test를 진행해 보았다.

귀무가설: 이산화질소의 여름 평균과 겨울 평균은 같다

대립가설: 이산화질소의 여름 평균이 겨울 평균보다 작다(단측검정)



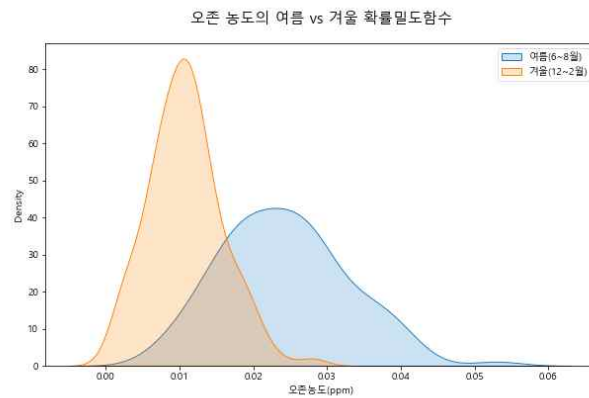
t-test 결과 p-value는 $7.905451383295384e-21$ 로, 0.05보다 작으므로 귀무가설을 기각할 수 있다. 따라서 이산화질소의 여름 평균이 겨울 평균에 비해 낮다고 말할 수 있고, 앞서 예측했던 계절성이 존재함을 통계적으로 확인할 수 있다.

6) 오존 계절성 검증을 위한 Two sample t-test

앞서 2018~2022 최근 5개년 월별 시계열 그래프 및 월별 평균 그래프를 통해 오존은 겨울에 감소, 여름에 증가하는 계절성이 있는 것 같다는 가설을 이끌어낼 수 있었으므로, 이를 통계적으로 검증하고자 한다. 이에 전체 데이터(모집단)에서 6~8월 범위의 여름 표본, 12~2월 범위의 겨울 표본을 추출해 Two sample t-test를 진행해 보았다.

귀무가설: 오존의 여름 평균과 겨울 평균은 같다

대립가설: 오존의 여름 평균이 겨울 평균보다 크다(단측검정)



t-test 결과 p-value는 $1.131882536536111e-32$ 로, 0.05보다 작으므로 귀무가설을 기각할 수 있다. 따라서 오존의 여름 평균이 겨울 평균에 비해 높다고 말할 수 있고, 앞서 예측했던 계절성이 존재함을 통계적으로 확인할 수 있다.

3. 기간별 일평균 데이터를 통해 요일 경향성을 살펴보자!

분석에 사용한 데이터: 서울시 기간별 일평균 대기환경 정보.csv

결측치: 없음

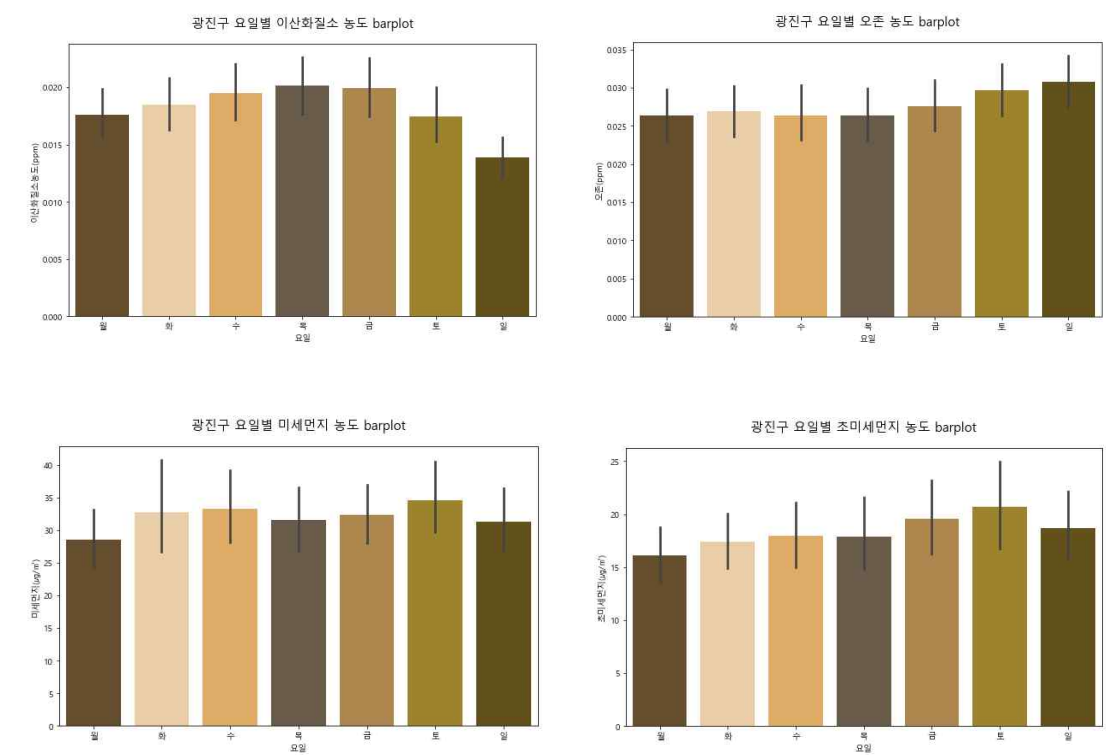
관측 기간: 2022.01.14.~2023.01.14

1) 각 대기오염물질별로 요일별 평균을 구한 결과

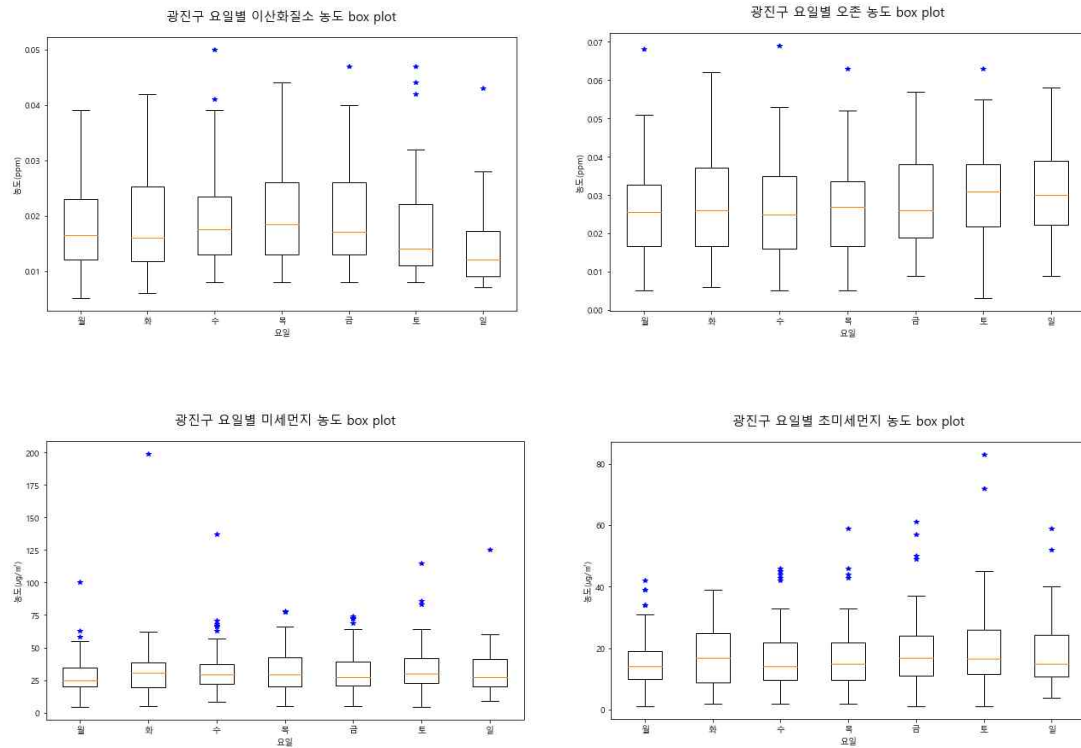
	최고 요일	최저 요일
이산화질소	목요일	일요일
오존	일요일	월요일
미세먼지	토요일	월요일
초미세먼지	금요일	월요일

오존, 미세먼지, 초미세먼지 세 대기오염물질이 최저 농도를 보이는 월요일을 피크닉 요일로 추천해도 좋을 것 같다.

2) 요일별 barplot



3) 요일별 box plot

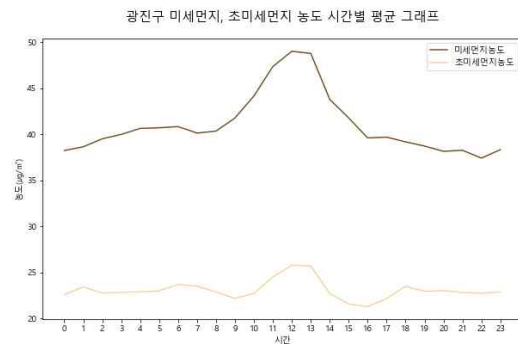
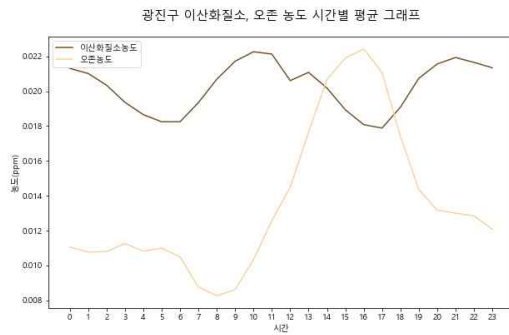


4. 기간별 시간평균 데이터를 통해 일변화 경향성을 살펴보자!

분석에 사용한 데이터: 서울시 기간별 시간평균 대기환경 정보.csv

결측치: 없음

관측 기간: 2022.11.14. 00:00~2023.01.14. 00:00



이산화질소: 8~10시, 20~22시의 출퇴근 시간대에 높다. 이는 출퇴근 시간대에 차량 이용량이 많아 이산화질소의 주요 배출원인 자동차 배기가스가 많아졌기 때문으로 추측된다.

오존: 14~17시의 오후 시간대에 높다. 오존은 앞서 이산화질소처럼 자동차 등의 배출원에서 직접 배출되는 1차 오염물질이 아니라, 질소산화물이나 휘발성유기화합물 등이 대기 중에 존재하는 상태에서 햇빛에 의한 광화학반응으로 생성되는 2차 오염물질이다. 오존 생성의 4대 요소로 NOx(질소산화물), VOCs(휘발성유기화합물), 자외선, 일정 이상의 온도를 꼽는다. 오존의 일변화를 봤을 때 낮 시간대에 최고치를 찍는 이유가 여기에 있다.²⁾

미세먼지, 초미세먼지: 11~13시를 정점으로 하고, 이를 전후로 한 아침, 저녁까지의 활동 시간대에 높고, 밤에는 비교적 낮은 양상을 보인다. 인간 활동에 의해 발생한 미세먼지가 큰 영향을 끼치고 있는 것으로 보인다.

5. 벚꽃 개화 시기에 유의미한 차이를 보이는 대기오염물질이 있을까?

분석에 사용한 데이터: 서울시 월별 평균 대기오염도 정보.csv

결측치

- 이산화질소 8개
- 오존 8개
- 미세먼지 162개
- 초미세먼지 308개

관측 기간

- 이산화질소, 오존: 1987.01~2023.01
- 미세먼지: 1995.01~2023.01

2)

<https://www.me.go.kr/mamo/file/readDownloadFile.do?sessionId=mGwdnrFkBMHnYTK4EMnUiL1M.mehome1?fileId=52279&fileSeq=1>

- 초미세먼지: 2012.01~2023.01

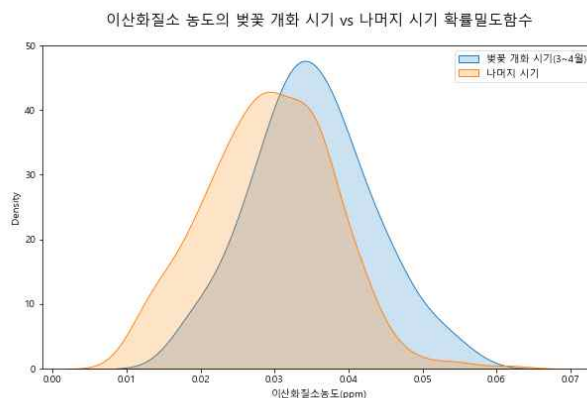
이번에는 t-test를 활용해 벚꽃 개화 시기(3~4월)에 유의미한 농도 차이를 보이는 대기오염 물질이 있는지를 검정하는 것이 목표이므로 따로 결측치 처리를 하지 않고, 각각의 대기오염 물질에 대해 t-test를 진행할 때 해당 대기오염물질의 결측치 행을 제거하는 방식으로 진행하였다.

이번 t-test의 목적은 사람들이 주로 한강 피크닉을 가는 벚꽃 개화 시기에 특정 대기오염물질의 농도가 다른 시기에 비해 높은지를 검정함으로써, 실제로 농도가 높다는 결론이 나올 경우 각별히 주의가 필요하다는 안내를 하기 위함이므로 특정 대기오염물질 농도의 벚꽃 개화 시기 평균이 나머지 시기 평균에 비해 높다는 방향으로 단측검정을 수행하였다.

1) 이산화질소 농도 차이 검증을 위한 Two sample t-test

귀무가설: 이산화질소 농도의 벚꽃 개화 시기 평균과 나머지 시기 평균은 같다

대립가설: 이산화질소 농도의 벚꽃 개화 시기 평균이 나머지 시기 평균에 비해 높다(단측검정)

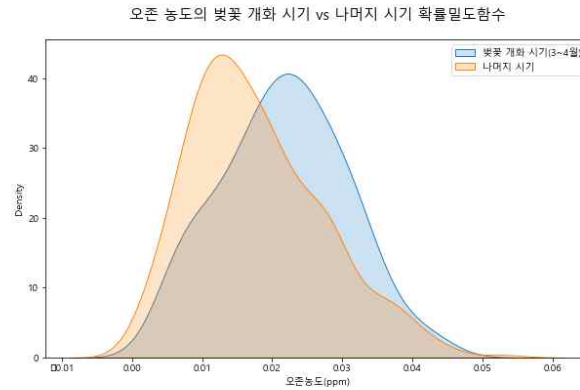


t-test 결과 p-value는 2.2244972366574967e-07로, 0.05보다 작으므로 귀무가설을 기각할 수 있다. 따라서 이산화질소 농도의 벚꽃 개화 시기 평균이 나머지 시기 평균에 비해 높다고 말할 수 있다.

2) 오존 농도 차이 검증을 위한 Two sample t-test

귀무가설: 오존 농도의 벚꽃 개화 시기 평균과 나머지 시기 평균은 같다

대립가설: 오존 농도의 벚꽃 개화 시기 평균이 나머지 시기 평균에 비해 높다(단측검정)

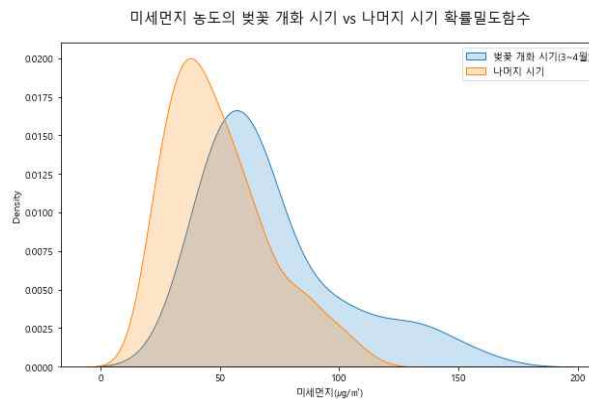


t-test 결과 p-value는 0.0012112292821281282로, 0.05보다 작으므로 귀무가설을 기각할 수 있다. 따라서 오존 농도의 벚꽃 개화 시기 평균이 나머지 시기 평균에 비해 높다고 말할 수 있다.

3) 미세먼지 농도 차이 검증을 위한 Two sample t-test

귀무가설: 미세먼지 농도의 벚꽃 개화 시기 평균과 나머지 시기 평균은 같다

대립가설: 미세먼지 농도의 벚꽃 개화 시기 평균이 나머지 시기 평균에 비해 높다(단측검정)

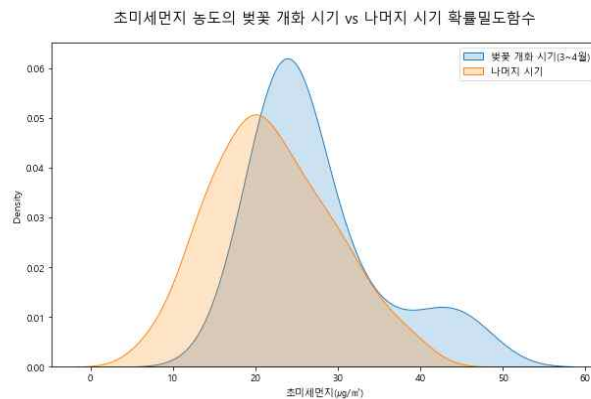


t-test 결과 p-value는 4.7005427852389025e-09로, 0.05보다 작으므로 귀무가설을 기각할 수 있다. 따라서 미세먼지 농도의 벚꽃 개화 시기 평균이 나머지 시기 평균에 비해 높다고 말할 수 있다.

4) 초미세먼지 농도 차이 검증을 위한 Two sample t-test

귀무가설: 초미세먼지 농도의 벚꽃 개화 시기 평균과 나머지 시기 평균은 같다

대립가설: 초미세먼지 농도의 벚꽃 개화 시기 평균이 나머지 시기 평균에 비해 높다(단측검정)



t-test 결과 p-value는 0.0018062700141789245로, 0.05보다 작으므로 귀무가설을 기각할 수 있다. 따라서 초미세먼지 농도의 벚꽃 개화 시기 평균이 나머지 시기 평균에 비해 높다고 말할 수 있다.

* 벚꽃 개화 시기 t-test에서 생각해봐야 할 점

월별 평균 그래프를 그려봤을 때, 미세먼지와 초미세먼지에 대해서는 확실히 벚꽃 개화 시기인 봄철(3~4월)에 다른 월에 비해 농도가 높은 경향을 확인할 수 있었기에 위에서 실시한 t-test가 의미를 가질 수 있으리라 생각한다.

하지만 이산화질소와 오존의 경우, 이산화질소는 겨울에 높고 여름에 낮으며, 오존은 겨울에 낮고 여름에 높은, 벚꽃 개화 시기와는 엇나가는 계절성을 가지고 있었다. 벚꽃 개화 시기와 나머지 시기를 비교하는 t-test를 진행해봤을 때 이산화질소와 오존 모두 벚꽃 개화 시기가 나머지 시기에 비해 농도가 높다고 할 수 있다는 결과를 얻을 수 있었으나, 이것이 정말 의미 있는 결과인지는 고민해봐야 한다. 이산화질소의 경우 여름철 농도가 낮은 것이, 오존의 경우 겨울철 농도가 낮은 것이 크게 작용하여 벚꽃 개화 시기의 농도가 상대적으로 크다고 나왔을 가능성이 있기 때문이다.