

한국 미세먼지 분석

이주영

분석 이유

어릴 때와 지금 생활의 가장 큰 차이가 있다면 초미세먼지라고 말할 수 있을 것이다.

어릴 때는 초미세먼지라는 말이 없었지만 요즘에는 오늘의 미세먼지를 체크하는 것이 하루 일과가 되었고, 초미세먼지가 심한 날 어두운 하늘을 보면, 환기를 시키기도 밖에 나가기도 두려웠다.

대기환경 데이터를 분석하여 지역별, 시간별 차이를 살펴보고 이에 대한 패턴과 인사이트를 찾고자 한다.

데이터

출처 **air korea**

데이터 총 용량 : **1.91GB**

사용 언어: **Python**

각 파일은 연도-월별 혹은 연도-분기별로 되어있었고

각 파일의 칼럼명은 지역, 망, 측정소명, 측정소코드, 측정일시, **SO2, CO, O3, NO2, PM10, PM25**로 되어있었다.

하지만 초미세먼지 (**PM2.5**)는 **2016**년부터 측정값이 존재한다.

지역, 측정일시, **PM10**(미세먼지 지수), **PM25**(초미세먼지 지수)만 골라 연도별로 새로운 데이터를 만들었다. 이 때, 각 파일의 미세먼지 지수들에 결측값이 있어 이들을 평균값으로 대체하였다.

데이터 탐색

미세먼지 데이터에 대한 질문들은 다음과 같다.

Q1. 미세먼지가 많아지면 초미세먼지도 많아지는가?

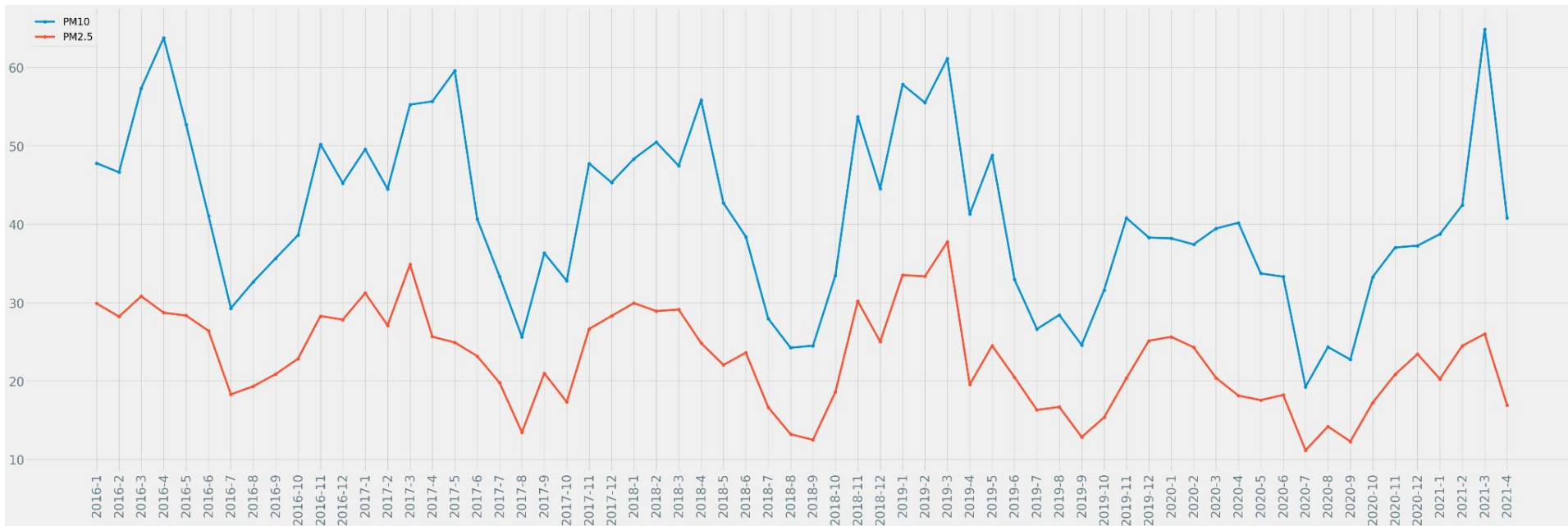
Q2. (초)미세먼지는 시간이 흐를수록 많아지는 추세인가?

Q3. 일년 중 미세먼지가 많아지는 특정 때가 있는가? 패턴이 있는가?

Q4. (초)미세먼지가 가장 많은/적은 지역은 어디인가?

Q5. 중국과 가까울수록 미세먼지가 많은가?

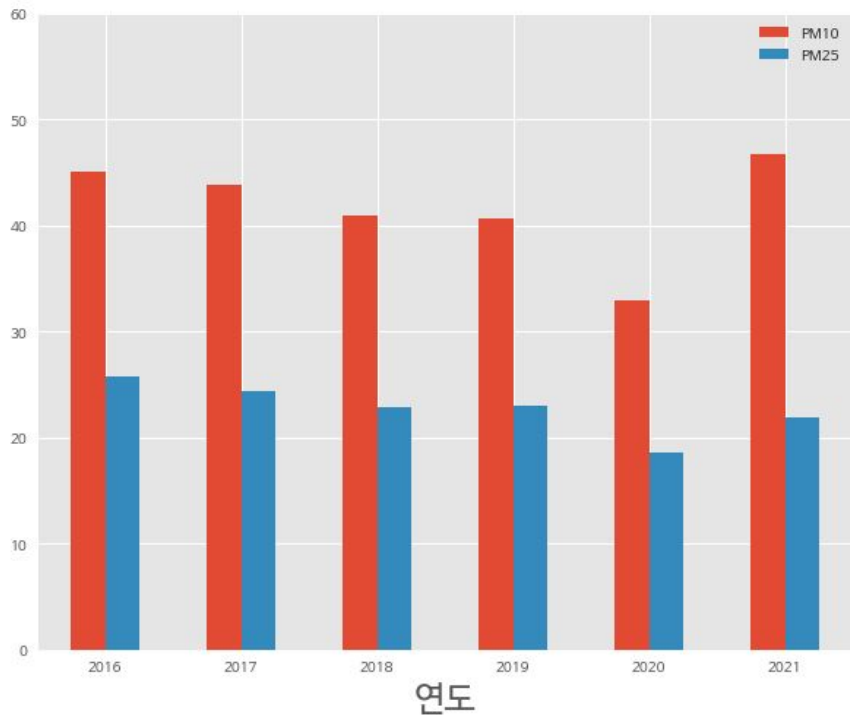
Q1. 미세먼지가 많아지면 초미세먼지도 많아지는가?



연도-월별로 **미세먼지**와 **초미세먼지** 지수를 시각화하였다.

미세먼지와 초미세먼지를 비교하니, 미세먼지가 감소하거나 증가하면, 초미세먼지도 증가하거나 감소하였다.

Q2. (초)미세먼지는 시간이 흐를수록 많아지는 추세인가?



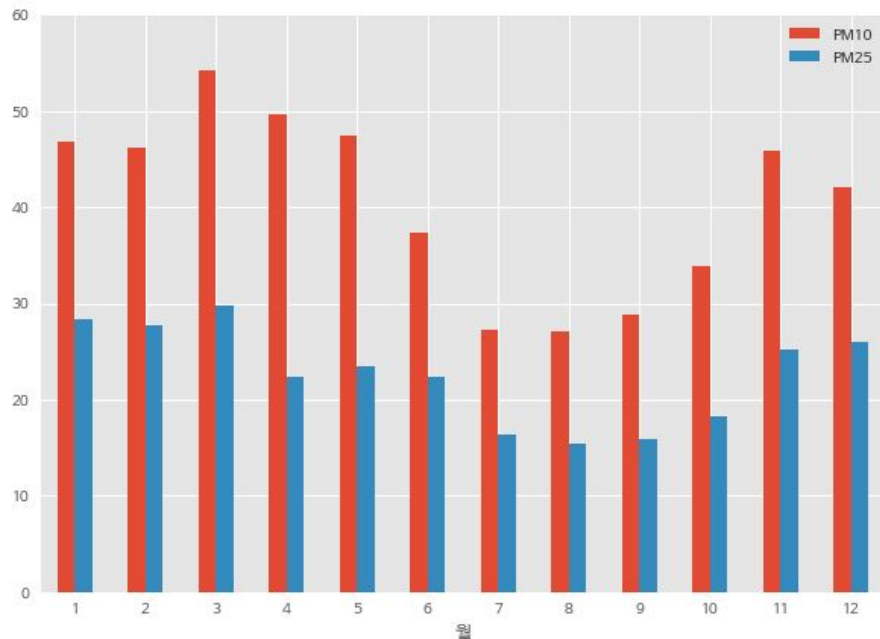
그렇지 않다.

2016년에서 2019년까지 감소세를 보이다가 2020년에서 큰 감소폭을 보이고, 2021년에 크게 증가하였다.

하지만, 2021년은 4월까지밖에 데이터가 없었기 때문에 코로나 회복이 시작된 2021년부터는 추이를 지켜볼 필요가 있다.

이는 아마도 코로나19의 확산이라는 변수가 원인으로 작용한 것으로 예상된다.

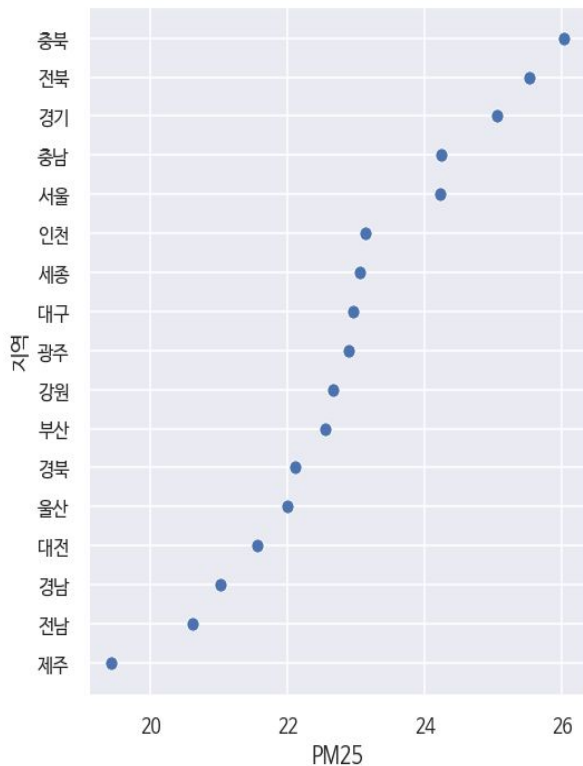
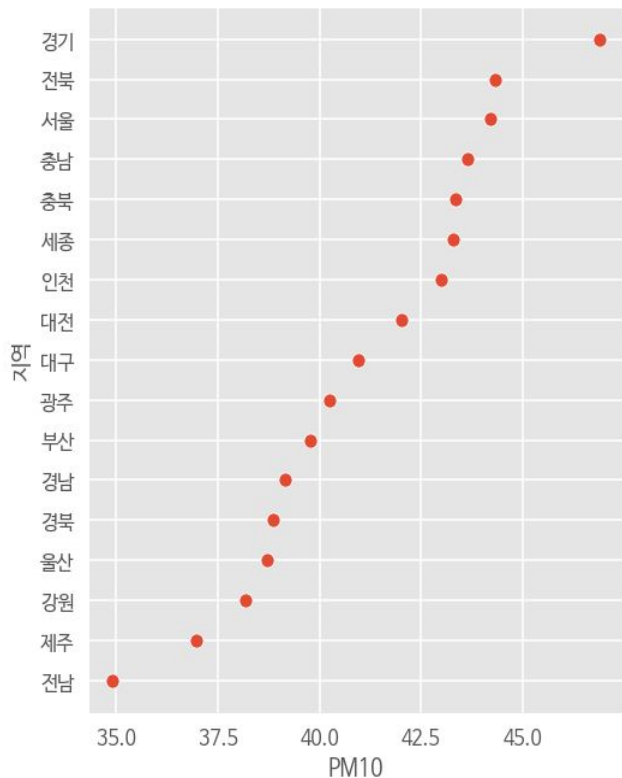
Q3. 일년 중 미세먼지가 많아지는 특정 때가 있는가? 패턴이 있는가?



월별로 살펴보았을 때 7월~9월이 가장 적었고, 3월, 4월이 가장 높았다.

이는 주로 3~5월에 발생하는 황사가 원인일 것으로 추정된다.

Q4. (초)미세먼지가 가장 많은/적은 지역은 어디인가?

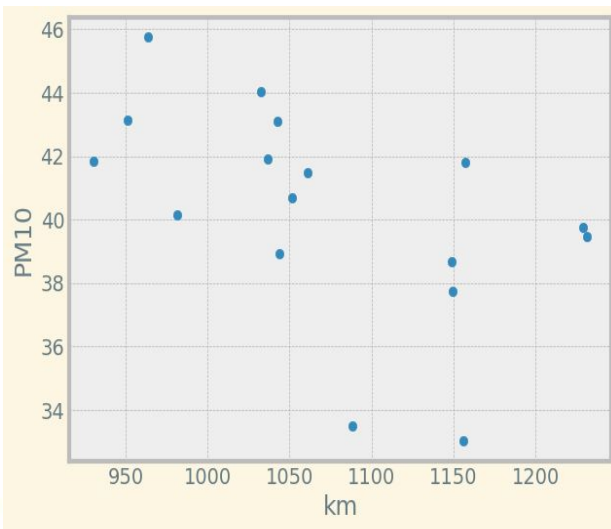
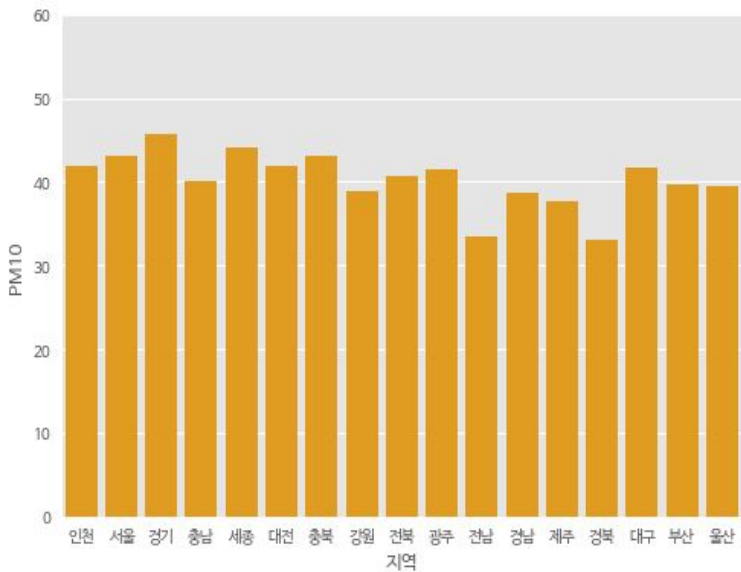


지역별로 미세먼지 지수가 큰 순으로 정렬한 결과, 미세먼지가 가장 많은 지역은 경기도, 가장 적은 지역은 전남이었다.

초미세먼지가 가장 많은 지역은 충북, 가장 적은 지역은 제주였다.

대체로 비슷한 순서였지만 강원도의 초미세먼지가 의외로 높았고, 전북 또한 예상보다 높은 결과가 나왔다. 하지만 대체로 수도권에서 높은 미세먼지가 관측되었다

Q5. 중국과 가까울수록 미세먼지가 많은가?



각 지역별로 중국과의 거리 데이터를 추가하고, **2013년부터 2021년** 미세먼지 데이터로 분석하였다

중국과 가장 가까운 지역 순으로 정렬하여 시각화한 결과, 거리별로 엄청난 차이를 보이지는 않았다.

거리와 미세먼지의 상관관계도 나타나지 않았다.

결론

미세먼지는 팬데믹이 시작된 **2020**년에는 현저히 줄고, 종식되어감에 따라 다시 크게 증가하는 모습을 보여주었다.

2021년은 4월까지밖에 데이터가 없었기 때문에 코로나 회복이 시작된 **2021**년부터는 추이를 지켜볼 필요가 있다.

대체로 수도권에서 높은 미세먼지가 관측되었지만 강원도의 초미세먼지가 의외로 높았고, 전북 또한 예상보다 높은 결과가 나왔다.

중국과의 거리와 미세먼지는 상관관계가 없었고, 미세먼지는 중국과의 물리적 거리만이 원인이 아니라 다양하고 복합적인 원인에 의해 나타난다.

그렇다면 미세먼지의 원인은 무엇일까?

석유, 석탄과 같은 화석연료 혹은 자동차 매연으로 인한 배출 가스에서 나오는 대기오염물질이 원인이라고 한다. 이것이 바람을 타고 주변 국가들에게도 영향을 미친다고 한다.

풍향 데이터와 국내 석탄화력, 경유 소비량 등의 데이터를 추가하여 분석하는 것도 좋은 방법이 될 수 있을 것이다.

감사합니다.