Air Pollution and Asthma From 2010-2021

Introduction:

Asthma is a large problem around the world, especially in the USA where 1 in 12 people have the condition (1). The severity of asthma is largely genetic there are a variety of factors that can cause asthma including the environment.

Air quality is a large contributor to allergy and lung inflammation which leads to asthma (2). Air quality is affected by a variety of things, the largest contributor being fossil fuels. Fossil fuels emit many harmful pollutants that stay in the air for a long time, including ozone, nitrogen dioxide (NO2), and small particulate matter (PM10) that will be discussed. In addition, natural disasters such as wildfires and earthquakes can cause an increase in PM10. These pollutants cause inflammations that can lead to asthma attacks and increase the intensity and number of asthma attacks.

Since the United States is large, there are many different lifestyles, climates, and topographies. This makes some states more suspectable to more asthma-prone conditions. Some states have more natural disasters than others such as California and Texas. California has wildfires and earthquakes while Texas is known for its tornados. In addition, the west coast and center of the United States are much dryer compared to other places which may indicate there is more particulate matter in the air. The urban cities also have more pollution due to cars and industries working there.

Considering all these factors, the purpose of this exploratory project is to see if there are any similar trends between pollutants and asthma prevalence between states.

Methodology:

The data for the pollutants was obtained using the Air Quality System (AQS) API from the Environmental Protection Agency (EPA). The R package used to extract the data was 'RAQSAPI' which allowed you to gather data based on selected criteria. Three pollutants were selected including ozone (code:44201), NO2 (coded: 42602), and PM10 (coded: 81102). The yearly average for each of the pollutants for each state was gathered using a loop.

The data for the asthma data was obtained from Behavioral Risk Factor Surveillance System (BRFSS) Prevalence Data collected by the CDC through web scrapping. The R package used to extract the data was 'xml2' and 'rvest' along with the help of 'tidyverse'. The asthma number and prevalence data were collected by the state for each year. A loop was used to collect the majority of the data.

Using the R packages, 'dplyr', 'tidyverse', and 'tidyr', the data was cleaned for only the needed columns such as the parameter, state, and pollution quantity for the AQS data. The data was merged and made wider by parameter.

To visualize the data, the R package 'latticeExtra' was used. This package allows to plot a double y-axis graph for each of the states. This package was used in the shiny object so that one can choose what state and pollutant trends they wish to see. The shiny object was created using the R package 'shiny'. This helps visualize what state and pollutant someone might want to see instead of having all the graphs all at once.

Link: https://luzmaria.shinyapps.io/Surv727/

Analysis:

To understand if there are any trends between the pollutant and asthma prevalence, as mentioned before, a double-axis plot was constructed for each state and pollutant. The most significant similarities for the pollutant ozone were states with high populations such as New York, Tennessee, Georgia, and Illinois as seen in Figure 1. These states have higher populations which indicates a higher need for fossil fuels, which is where ozone comes from. A majority of the states did not have any significance, in actuality, a majority of rural states like Maine, Alaska, and Montana, had no real changes in trends , as seen in Figure 2. The ozone level for the more rural states was much lower than for the more populous states. Overall, there wasn't any significance in trends between ozone and asthma prevalence.
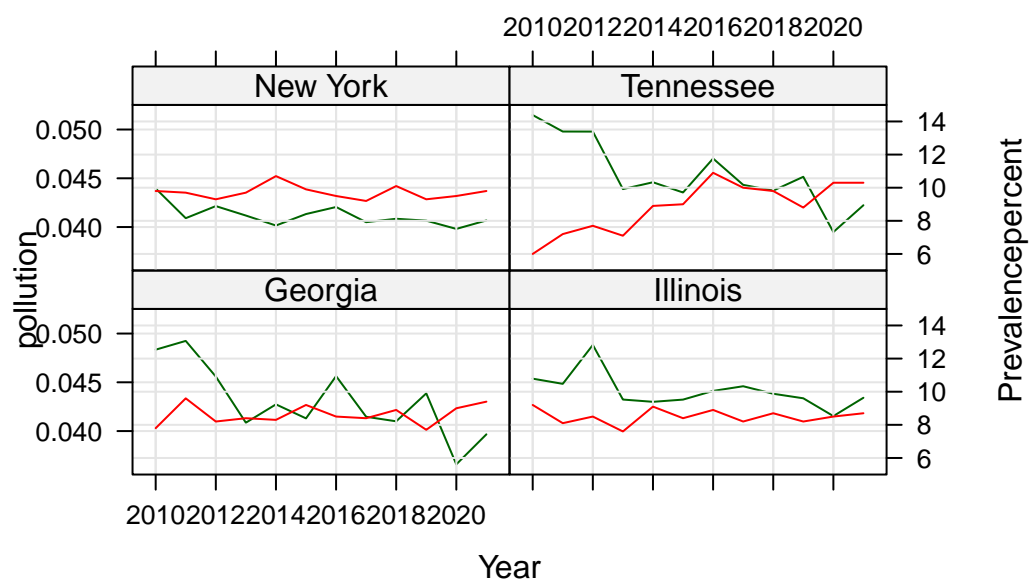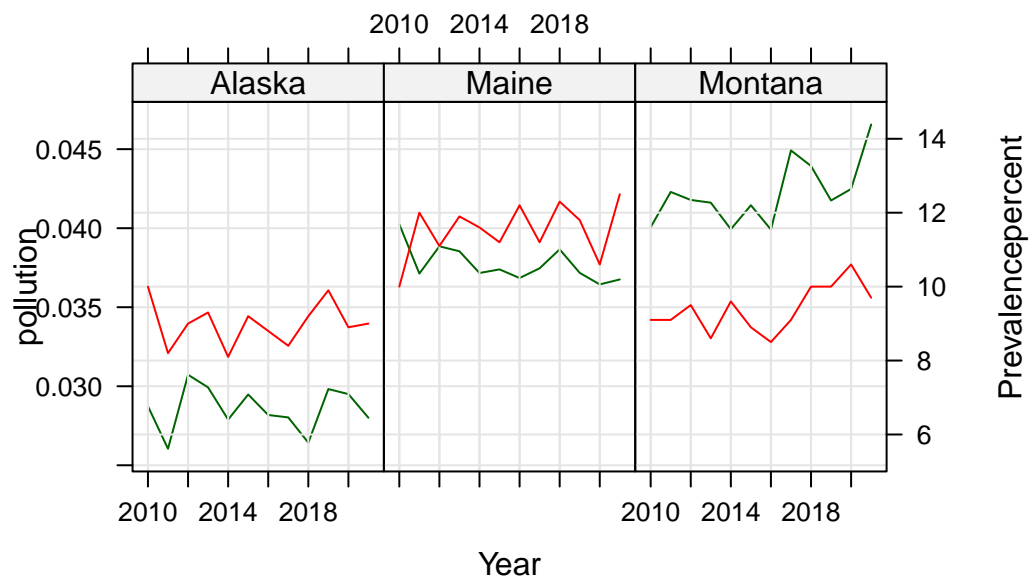
Figure 1: The similar trends with Ozone.

Figure 2: The different trends with Ozone.

PM10 was also visualized. PM10 was the pollutant of most interest due to the more direct

impact on lung inflammation thus leading to higher asthma attacks which could in theory lead to higher asthma prevalence. The most significant similarities between the PM10 trend and the asthma prevalence trend are in Nevada, California, Texas, New Mexico, New Jersey, and Michigan as shown in Figure 3. These states are high in population and have high industrial facilities which correlates to a higher PM10 exposure. In addition, many of the states are known for their dry climates such as Nevada and California, a dryer climate may produce a lot more dust which increases PM10 in the air. Many of the less populous states such as South Dakota, Vermont, North Dakota, and Maine (Figure 4) do not have any real trends in their PM10 trends since they do not have much use for fossil fuels and are in more wetter climates. Overall, PM10 had interesting trends with asthma prevalence.
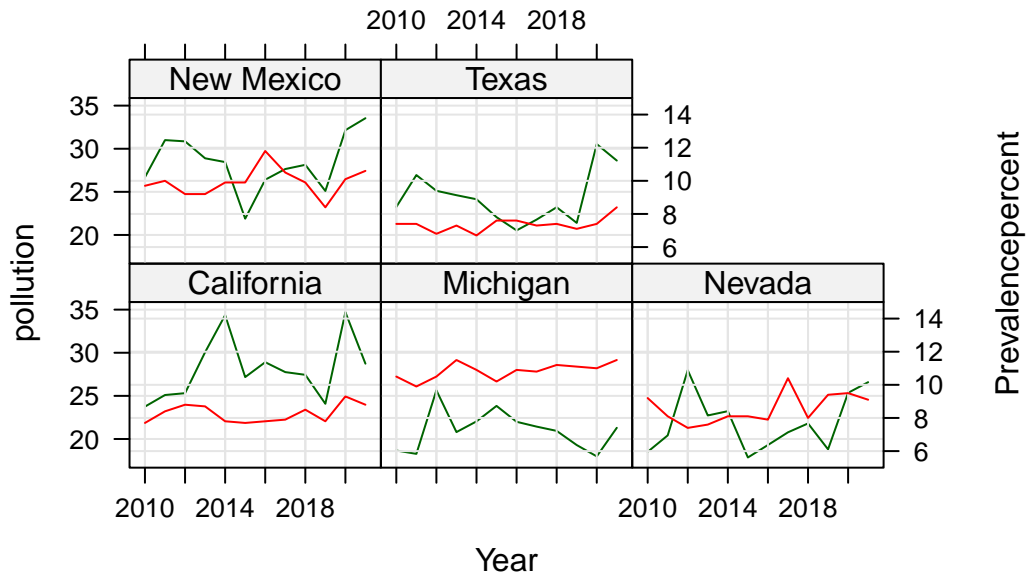

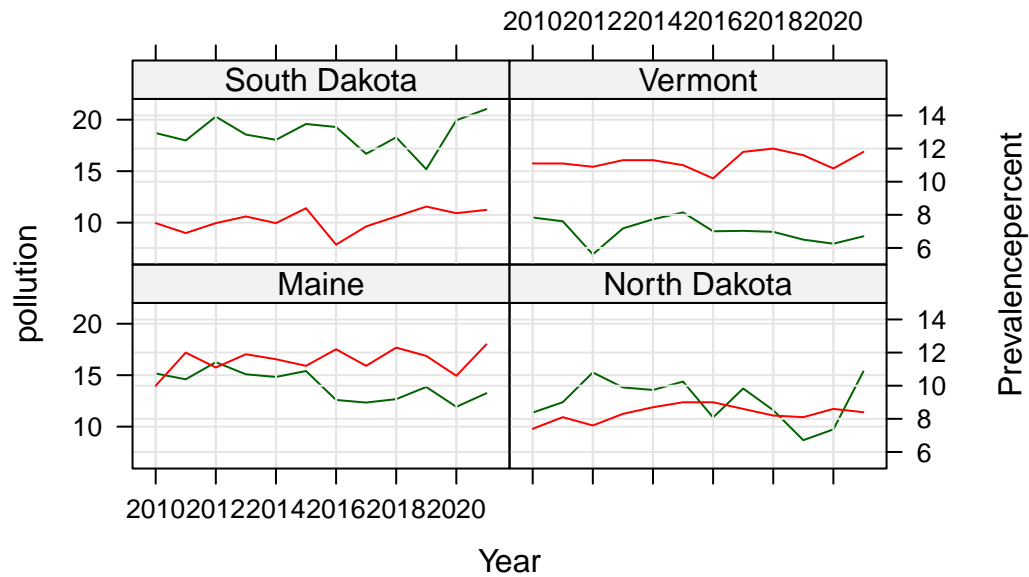Figure 3: The similar trends with PM10.

Figure 4: The different trends with PM10.

Nitrogen dioxide was also visualized using the double y-axis. NO2 is also interesting because it is such a large by-product of fuel fossil consumption and power plants, so it is expected to be quite high in largely urban and industrial facilities. The most significant similarities between the trend of NO2 and asthma prevalence are in the states, of Nevada, California, New York, Illinois, and Ohio as shown in Figure 5. These states again, are high in population (many major cities) and have a lot of industrial facilities. It makes sense for these states to have higher NO2 levels due to their higher consumption of fossil fuels. The states that had less trends are states that have low activity such as Louisiana, Kansas, and Utah (Figure 6), along with others. There is no real trend between them. Overall, the visual data from NO2 is similar to that of PM10.
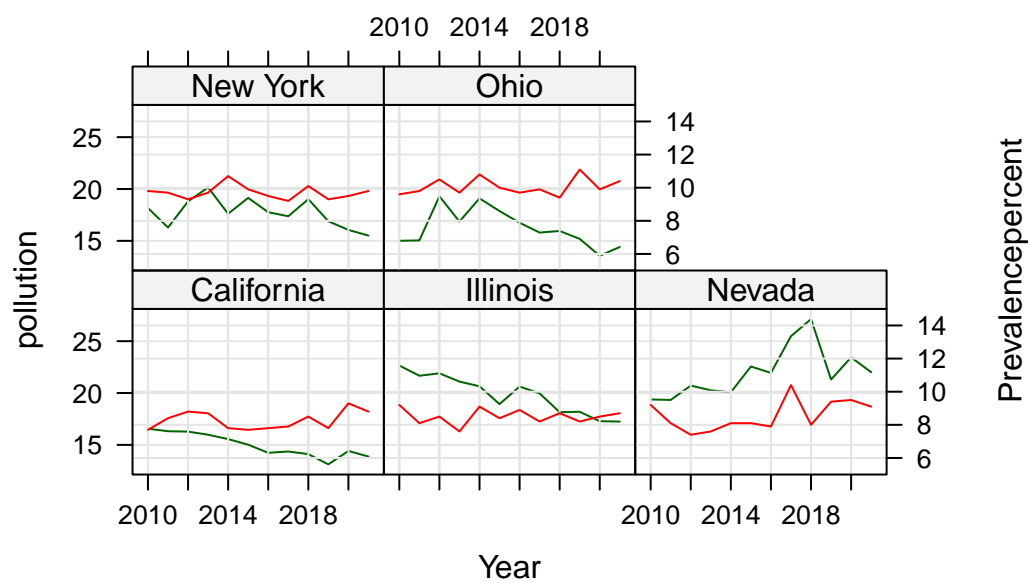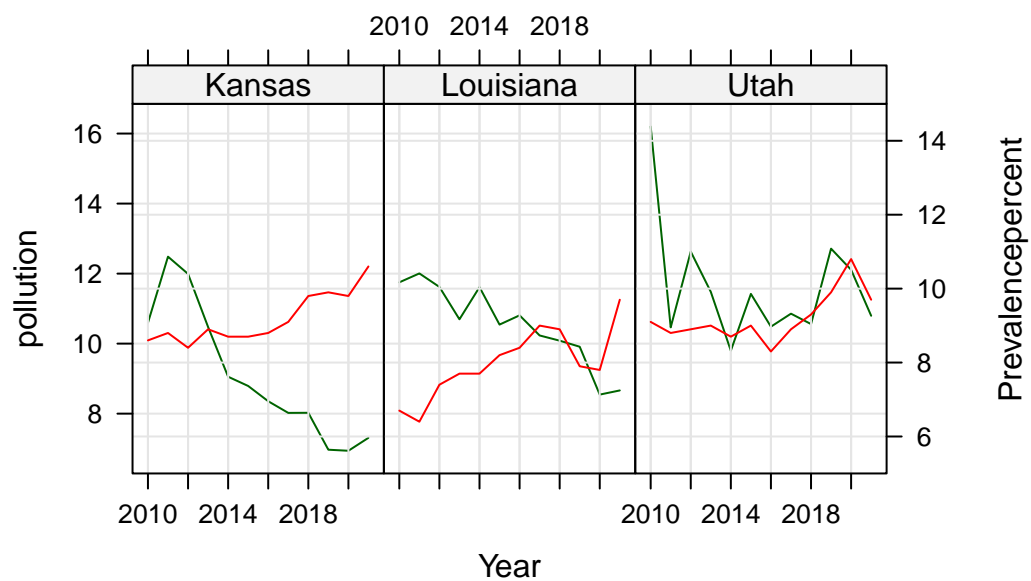
Figure 5: The similar trends with NO2.

Figure 6: The different trends with NO2.

6

Conclusions:

This exploratory project had some interesting findings. According to the visual representations, ozone was not a major contributor to asthma prevalence. It may be that the concentration is much too low to induce lung inflammation. Even in locations where there is a high population and high fossil fuel activity, there was not a clear trend in the asthma prevalence with ozone. In terms of PM10 and NO2, there were more trends that could be followed especially for the states that have higher populations and fossil fuel activity. These states followed trends more closely.

The way I chose to determine if the lines were closely related is by looking at the trends and seeing if the highs and lows aligned because it is expected that with higher pollution rates, there would be higher asthma prevalence.

I wanted to add a more qualitative measurement of similarity and I thought a correlation between them might work but it could not be added since the trends are not linear. Since the trend isn't linear, the correlations would not say anything significant. I added the code below to the supplement.

The limitation of this project is that this is all visualization data, this is not the complete answer to the exploration question. There are many future questions to be answered such as if there is a difference in child and adult asthma prevalence or to look at a closer time frame like monthly instead of yearly. Since the data is yearly there may have been a difference in asthma prevalence by month. In addition, it would also be interesting to look at asthma attacks and see if there are any trends there.

Supplement:

In addition to further analysis, the correlation between all the states and pollutants was done using a loop and added to a table.

```
# A tibble: 45 x 4
   State         PM10     NO2   Ozone
   <chr>        <dbl>   <dbl>   <dbl>
 1 Alabama     -0.318 -0.0419 -0.0365
 2 Alaska      -0.318 -0.0419 -0.0365
 3 Arizona     -0.318 -0.0419 -0.0365
 4 Arkansas    -0.318 -0.0419 -0.0365
 5 California  -0.318 -0.0419 -0.0365
 6 Colorado    -0.318 -0.0419 -0.0365
 7 Connecticut -0.318 -0.0419 -0.0365
 8 Delaware    -0.318 -0.0419 -0.0365
 9 Florida     -0.318 -0.0419 -0.0365
```

```
10 Georgia     -0.318 -0.0419 -0.0365
# i 35 more rows
```

GitHub: https://github.com/luzmariav/FinalProject

References:

1. Asthma and Allergy Foundation of America.   https://aafa.org/asthma/asthma-facts/#:~:text=How%20Common%20Is%20Asthma%3F,about%201%20in%2012%20people.&text=More%
2. American Lung Association. https://www.lung.org/lung-health-diseases/lung-disease-lookup/asthma/learn-about-asthma/what-causes-asthma