WildLand Fire Analysis for Indianapolis, IN

DATA 512: Human-Centered Data Science
Himanshu Naidu

1. Introduction

The capital and largest city of the U.S. state of Indiana, Indianapolis, also referred to as "Indy," serves as the seat of Marion County [1]. Indianapolis is located along the west fork of the White River in the central till plain region of the state.

Public health is becoming increasingly concerned about the frequency and severity of wildfires brought on by climate change all over the United States. Hazardous chemicals found in dense smoke from wildfires greatly increase the risk of respiratory illnesses, hospital stays, and early mortality [2] [3]. Vulnerable groups are disproportionately impacted by these health issues. Using a human-centered data science methodology, this analysis highlights the health outcomes of those affected by wildfires and investigates the relationship between exposure to smoke from these fires and respiratory health.

This project aims to answer important issues regarding the long-term impacts of smoke exposure on Indianapolis residents' respiratory health. The results are intended to serve as a guide for community organizations, public health professionals, and local officials as they adopt interventions and policies to lessen these effects. This study offers important insights into the public health responses and policy actions required to support affected communities by highlighting community readiness and resilience.

2. Background

2.1 Related Work

Due to the increased frequency and intensity of wildfires brought on by climate change, research on the effects of wildfire smoke on socioeconomic situations and public health has increased. Wildfire smoke can be extremely harmful to the lungs, especially for children, older adults and those with asthma, COPD and bronchitis or a chronic heart disease or diabetes [4].

Particle pollution, a mixture of extremely small solid and liquid particles hanging in the air, is one of the numerous contaminants included in wildfire smoke. Because they are so tiny, these particles can enter the lungs and become lodged there. Particle pollution can cause fatalities as well as heart attacks, strokes, and asthma attacks [5]. Individuals who breathed the smoky air during wildfires in California were more likely to cough, wheeze, get bronchitis, get colds, and need to visit the doctor or hospital for respiratory reasons, particularly asthma, according to studies on the subject [7].

Carbon monoxide (CO), a colorless, odorless gas that is most prevalent during a fire's smoldering stages and in the vicinity of the fire, is another hazard caused by smoke from forest fires. CO can cause a number of respiratory problems, headaches, nausea, dizziness, and, in high quantities, early death. It also decreases the amount of oxygen that reaches the body's organs and tissues. Numerous other dangerous emissions, such as nitrogen oxides and other dangerous air pollutants, are dispersed by wildfires.

Understanding the potential impacts of wildfire smoke on the residents of Indianapolis, is thus essential for informed decision-making by local authorities and residents.

2.2 Research Questions

This focus for this project is rooted in a human-centered approach, directly addressing the well-being and safety of Indianapolis residents. For this project, I have chosen to focus on the healthcare sector within Indianapolis. More specifically, I will be assessing the impact of wildfire smoke on asthma and COPD.

Using data on asthma and COPD fatalities and occurrences from 1999 to 2020, I verify the association between wildfire smoke and chronic respiratory diseases in Indianapolis. Using a customized smoke estimate, which takes into account the size, distance, and kind of fire from any wildfire within 650 miles of Indianapolis, IN, I estimate the aggregated impact of the wildland fires in proximity, in this analysis. After that, I project how wildfires will affect each health parameter over the following ten years.

In this project, I direct my focus on the following research questions:

1. How do the trends of wildfire smoke and respiratory illnesses correlate with each other?

For this question, the initial hypothesis is that the increasing smoke levels over the past decades could be associated with an increase in incidences of and deaths caused by respiratory illnesses.

2. Using smoke estimates as an independent variable, how are the trends of respiratory illnesses expected to change in the coming years?

For this question, the initial hypothesis is that with the hypothesized rise in wildfires, even the impact of respiratory illnesses will be worsened in the future.

2.3 Datasets Used

2.3.1 Wildfire Data

The cleaned and organized wildfire data used in this project was provided by the USGS [8]. The final geospatial dataset that was created contains both a raw, merged dataset with duplicates and a "combined" dataset that is duplicate-free. This combined dataset includes both wildfires and prescribed fires from the mid-1800s to 2021, compiled from 40 original wildfire datasets. For analysis, I utilized only the combined dataset that is provided as a Geodatabase file created on ArcGIS.

We focus on a select subset of Wildland fires that satisfy the following conditions:

1. Occurred in the last 60 years: [1964 - Present]:

This is straightforward to filter and is intended to avoid potentially inaccurate data collected prior to the advent of satellite imaging

2. Within 650 miles of Indianapolis, IN:

We designate (39.7684, -86.1580) as the latitude and longitude coordinates representing the center of Indianapolis.
For distance calculation, we go with EPSG:5070 (Conus Albers / USA Contiguous Albers Equal Area Conic). This CRS is indeed well-suited for the contiguous United States and designed specifically to minimize distortion for both area and distance across this region.

3. Occurred in the fire season (1st May - 31st October):

In the given Geodatabase file, we had the Listed_Fire_Dates attribute in the database that we utilized to get the possible start and end dates for each wildland fire, and filtered out those that did not have an overlap with the fire season.

2.3.2 Air Quality Index Data

This analysis required historical Air Quality Index (AQI) data for Indianapolis, IN, focusing on fire season (May 1 - October 31) each year from 1963 onward. The AQI measures daily air quality to indicate how healthy the air is, tracking pollutants like smog and smoke. Typically, an AQI between 0-50 reflects healthy, clean air, while 500 marks the upper limit for hazardous conditions.

For this project, I accessed data via the US Environmental Protection Agency (EPA) Air Quality Service (AQS) API [9].

2.3.3 Respiratory Health Data

The following sources of data are utilized to extract respiratory health data in Indianapolis, IN.

I. <u>Global Burden of Disease (GBD)</u>
The data contained in the dashboard provided by IHME, which stands for Institute for Health Metrics and Evaluation [10], can be used to extract respiratory illness-related raw counts and rates (per 10,000) for incidence (new cases), prevalence (total cases), and deaths, from 1999-2021. The lowest level of granularity is state-level (Indiana).

From the dashboard, we can get access to a CSV file that contains all the aggregated data to be obtained from the dashboard. This includes state-level numbers and rates data on Incidents and Deaths of respiratory illnesses for every year.
The datasets can be downloaded from the dashboard in zipped CSV les. The data is governed by the IHME FREE-OF-CHARGE NON-COMMERCIAL USER AGREEMENT [11].

II. MCPHD Environmental Public Health Tracking Dashboard
The data contained in this dashboard tracks asthma and COPD-related ED (Emergency Department) visits and hospitalizations from 2016 - 2023. The main reason for the limited window of data is the fact that Marion County was only recently awarded a grant by the CDC (Centers for Disease Control and Prevention), in 2022.

An important point to note is that due to the extremely limited data available currently in this data source, the Marion County related data was only used for Exploratory Data Analysis, and had to be dropped for the final analyses.

III. Federal Reserve Economic Database (FRED)
FRED is an online database that has more than 800,000 time series datasets from various sources. From this database, we will be utilizing the 'Age-adjusted premature death rate' for Marion County and Indiana. We can access the multiple csv files that contain all the aggregated data for the age-adjusted premature deaths in both Marion County as well as Indiana.

FRED explicitly mentions that all data can be freely utilized, as long as one mentions FRED as the service from which the data was retrieved from, and keep note of the copyright notices that appear on the data with FRED. Additional details for the terms of use can be found here.

2.4 Models Used

For this analysis, I used open-source time series models such as AutoRegressive Integrated Time Series (ARIMA), Vector AutoRegression (VAR), Vector Autoregressive Moving-Average with Exogenous variables (VARMAX), which are effective in forecasting both univariate and multivariate time series.

3. Methodology

3.1 Smoke Estimate Calculation

The aim is to provide an annual estimate of wildfire smoke levels in Indianapolis, Indiana, to better understand how wildfires affect city residents. Smoke levels are typically influenced by several factors such as wind patterns over days, fire intensity, duration, proximity to the city, etc. For this analysis, however, we're working with just the fire area and distance data.

I also use the fire type—distinguishing between prescribed burns and true wildfires—as a stand-in for fire intensity. Prescribed burns, usually planned for optimal weather conditions, are intended to reduce both safety risks and smoke spread, meaning they likely contribute less to smoke exposure in nearby cities compared to wildfires.

Taking inspiration from the "inverse-square law," which is common for concepts like light, sound, and gravity, we can postulate the relationship between smoke, distance and area to be the following:

$$\text{Smoke} \propto \frac{\text{Area}}{\text{Distance}^2}$$

As mentioned before, I hypothesize that true wildfires have a greater impact on smoke. We will be encoding fire-type as a quantitative variable, where the value of impact is lower if the fire is controlled.
I postulate the relationship between smoke and fire-type to be the following:

$$\text{Smoke} \propto Fire\_Type$$

(where $Fire\_Type_{true} > Fire\_Type_{prescribed}$)

I would also take into consideration some interaction between distance, area and fire-type. Taking all these relationships into account, we can have the following formula (based on the linear regression formula, and the □ are coefficients, that are tunable hyperparameters for our model):

$$\text{Smoke} = \beta_0 + \beta_1 \frac{\text{Area}}{\text{Distance}^2} + \beta_2 Fire\_Type + \beta_3 \frac{\text{Area}}{\text{Distance}^2} Fire\_Type$$

For the given model, we will be giving the betas the following values (thus, accounting only for interaction among all the variables):
$\square_0 = 0$, $\square_1 = 0$, $\square_2 = 0$, $\square_3 = 1$

Using the given wildfire-specific smoke estimates, I calculated annual smoke estimates for each year, using the smoke estimates for each wildland fire in that year, weighted by the number of days the fire was active in the fire season.

3.2 Comparing Smoke Estimate with AQI

I verified my smoke estimations by comparing them with the annual AQI data for Indianapolis. In particular, daily AQIs for both gaseous and particle pollutants for each day of the fire season were obtained from monitoring stations located within the city of Indianapolis. To determine the final yearly AQI values, the maximum AQI value for each day was selected, and then an average across the fire season was calculated.

3.3 Forecasting Smoke Estimate

Once I validate my smoke estimate, I then use an AutoRegressive Integrated Moving Average (ARIMA) [12] model for forecasting future smoke estimates for the next 25 years. ARIMA models are unable to adequately account for time-varying variance, even if they can handle non-stationarity to a certain extent. To determine if the mean and variance of our data are constant, I employed the Augmented Dickey-Fuller (ADF) test.

3.4 Smoke Estimate and Respiratory Illness Factors Correlation

Using our smoke estimates, and the state-level respiratory data retrieved from IHME, I explored the possible relations between variables of interest. At the state level, I have access to respiratory health indicators that show the death, incidence, and prevalence rates of COPD, asthma, and all other chronic respiratory illnesses expressed as rates per 100,000. Using Year as the primary key, I combine these diverse datasets with the wildfire data.

I examine the trends as well as the pairwise correlations between each statistic to determine any possible connections between them and my annual smoke estimates. In particular, I examine the Pearson correlation coefficients. It is assumed that there is a linear relationship between the two indicators and that the data is normally distributed with no outliers in order to perform the Pearson correlation coefficient tests.

3.5 Forecasting Respiratory Illness Factors

Using my smoke forecast and time as the exogenous factors, I model multivariate time series using a Vector Autoregressive Moving Average with eXogenous regressors model (VARMAX) in order to predict the respiratory health indicators. Since many of the health markers are naturally closely related, I employ a multivariate approach. We can estimate the impact of one of these indicators on other variables using VARMAX, in contrast to ARIMA. In particular, I make future predictions for groups of variables at a time that are all connected to the same disease. For instance, in order to predict asthma rates, I take into account both the incidence and the number of deaths, with smoke estimates as the exogenous variable. I predict COPD using the same related variables.

For each forecast, I divided the data into 80% train and 20% test before developing the model. We used the Euclidean norm of the Root Mean Squared Error to track model fit and performed a grid search for the optimal hyperparameter values of p & q.

4. Findings

4.1 Exploratory Data Analysis

At first, I performed some exploratory data analysis on the wildland fire data to check if I find some interesting patterns.

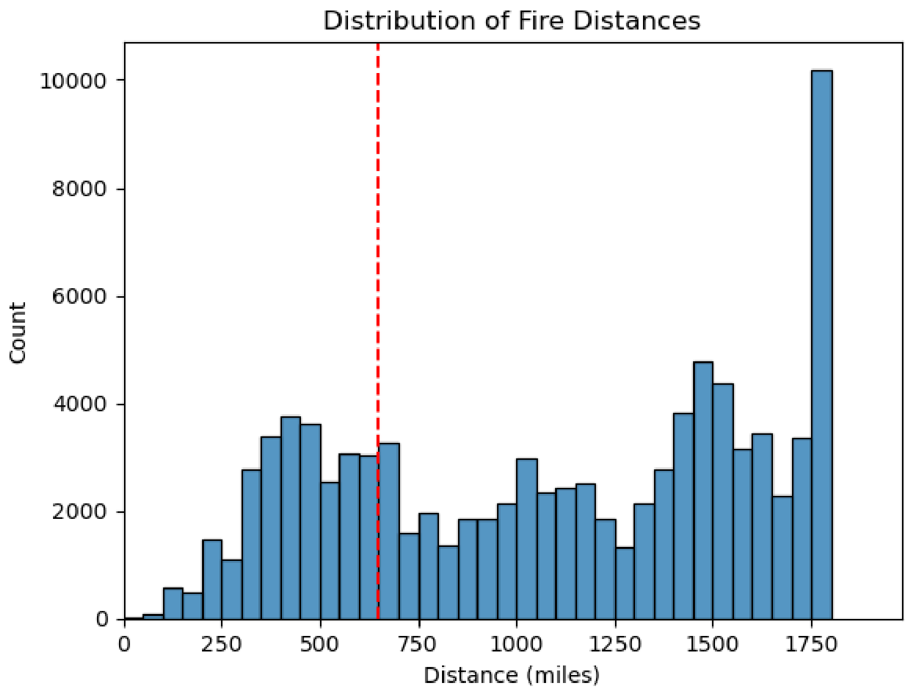

Figure 1: Histogram for Distribution of Fire Distances from Indianapolis

The histogram in Figure 1 shows the frequency of fires occurring at various distances from Indianapolis, IN, organized into 50-mile bins. The data comes from the combined USGS ArcGIS GeoDatabase, which includes attributes like fire date, distance from the city, smoke estimates, area burned, fire type, and additional details.
For this visualization, we used the number of fire occurrences within each 50-mile distance interval, up to 1800 miles. The x-axis displays the distance from Indianapolis in miles, and the y-axis shows the number of wildfires. For additional detail, we have also divided each bar in the chart by its distribution of fire type. The distribution appears fairly even, with a greater concentration of fires at further distances from the city.

The highest number of fires occurs around the 1800 mile distance from Indianapolis, suggesting that Indianapolis itself is not a major wildfire hotspot, as fewer fires are observed closer to the city. Finally, the dotted red line indicates the distance threshold (=650) that is used to perform the modeling and analysis.

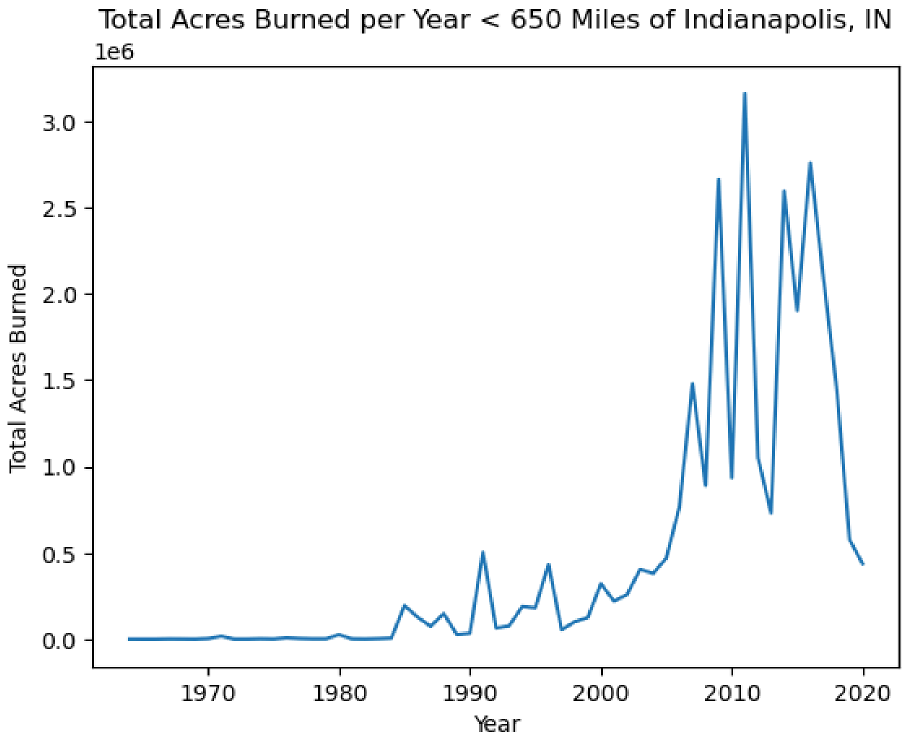

Figure 2: Time Series Graph of Total Acres Burned Per Year (within 650 miles of Indianapolis)

The time series graph in Figure 2 shows the area burned by fires near Indianapolis, IN. For this visualization, I used the yearly total area burned by fires. The x-axis represents the years from 1963 to 2020, and the y-axis displays the total burned area in units of 10^6 acres to improve readability. It's evident that the area affected by fires has been increasing over time, with a significantly higher burn impact in the 2000s compared to the 1900s. Since the late 2000s, the trends have fluctuated massively, but on average, been quite high. Overall, this suggests that the impact of fires around Indianapolis has grown substantially over the last few decades.

4.2 Smoke Estimates

On comparing the smoke estimates with AQI data, as seen in Figure 3, surprisingly, we see that the annual AQI values have gotten lower over time, in Indianapolis, while my smoke estimates keep increasing. In the initial years, the smoke estimate is quite low while the AQI values are relatively high. Over time, this trend reverses, with smoke estimates increasing and AQI values declining. This negative correlation suggests that, even as the prevalence and intensity of fires grew, Indianapolis's air quality improved, potentially due to effective air quality control measures implemented in Indiana. Consequently, despite an increase in smoke-generating fires, Indianapolis has managed to maintain or even enhance local air quality over the years.

Indeed, on further research of the state of Indiana as a whole (for a better overall picture), we see that there have been measures taken to improve the air quality in Indiana. For instance, the 2021 Fine Particles (PM2.5) [6]. Data Summary Report by IDEM shows a reduction in PM2.5

levels, which are tiny particles in the air that can pose health risks. This decline is attributed to regulatory measures and efforts to reduce emissions from various sources.
However, it must be noted that AQI values do not serve as a source of truth, since the metrics do not comprise only smoke particles. Thus, while they serve as an initial validation, we cannot use them to improve our smoke estimates. Hence, our analysis will continue with our existing smoke estimates.

To further validate these visible contrasting trends, I check the Spearman correlation, which turns out to be -0.38 with a p-value of 0.006, which indicates a weak negative correlation.

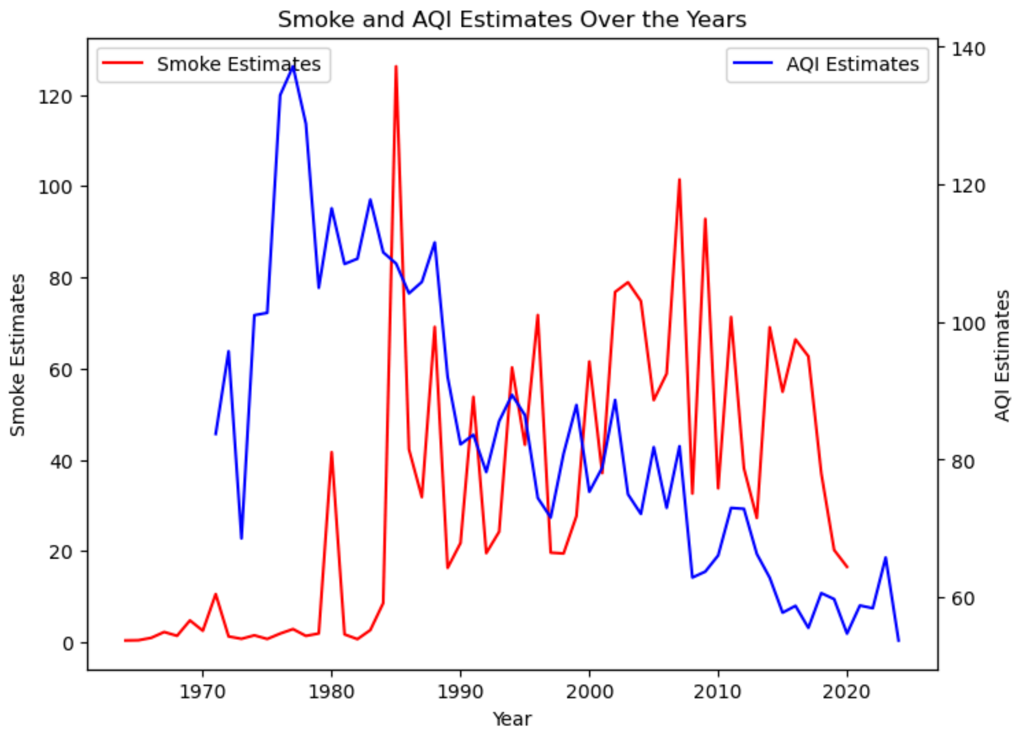

Figure 1: Comparing Smoke Estimates with Annual AQI Values

While forecasting my smoke estimates for the next 25 years, the ADF test value observed was -2.75, along with a p-value of 0.006, thus indicating with high confidence, that the distribution is stationary. Thus, using the smoke estimates data for the past 40 years, I forecasted the estimates for the next 25 years, to achieve the result as given in Figure 4.

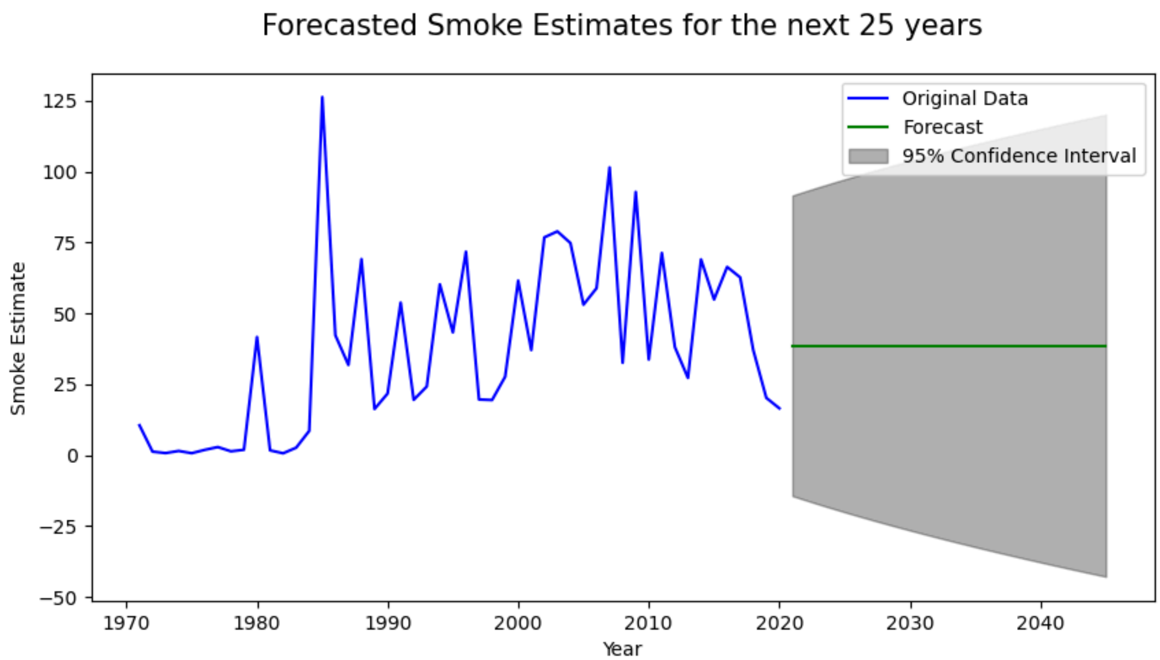

Figure 2: Forecasted Smoke Estimates for 2021-2045

Surprisingly enough, the model forecasts that the smoke estimates will be relatively constant for the next 25 years. While it may be an unusual result, there might be some confounding factors at play here that we unfortunately cannot verify with the absence of extra data such as wind patterns.

4.3 Smoke Estimates and Respiratory Illness Factors Correlation

The pairwise correlations among the smoke estimate and the various respiratory illness factors, are observed in Figure 5, and as one can see, there isn't a strong linear correlation between the respiratory illness factors and my smoke estimates, although this does not preclude the possibility of a non-linear correlation among the factors (analyzing the trends visually however, did not indicate any possible correlations).

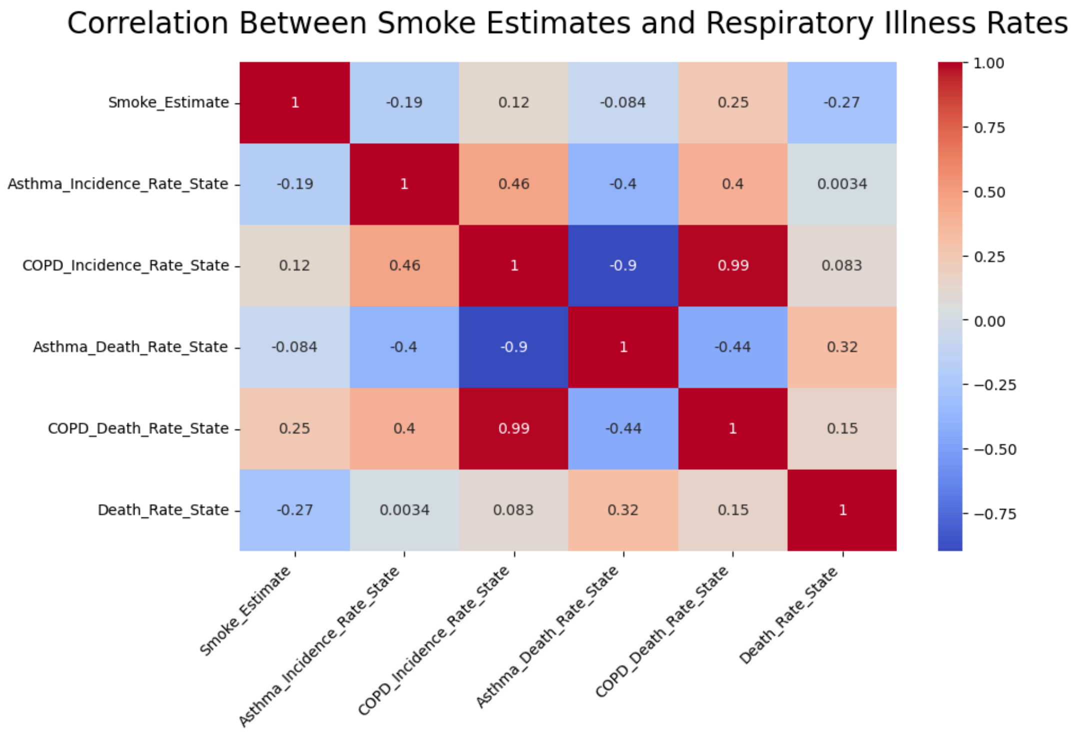

Figure 5. Correlation between Smoke Estimates and Respiratory Illness Factors (along with State-Level Death Rate)

If we look at the respiratory health indicators at the state level, we see that generally, they do not have a very significant correlation with smoke estimates, as shown in [Figure 6] and [Figure 7].

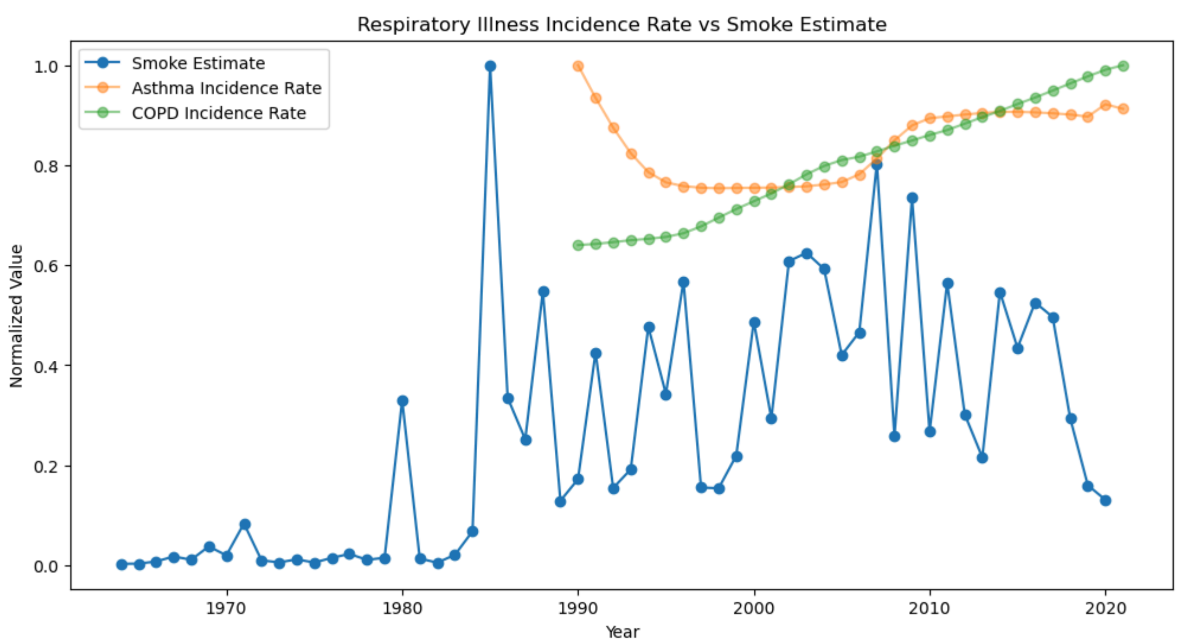

Figure 6: Trends of Respiratory Illness Incidence Rates vs Smoke Estimates

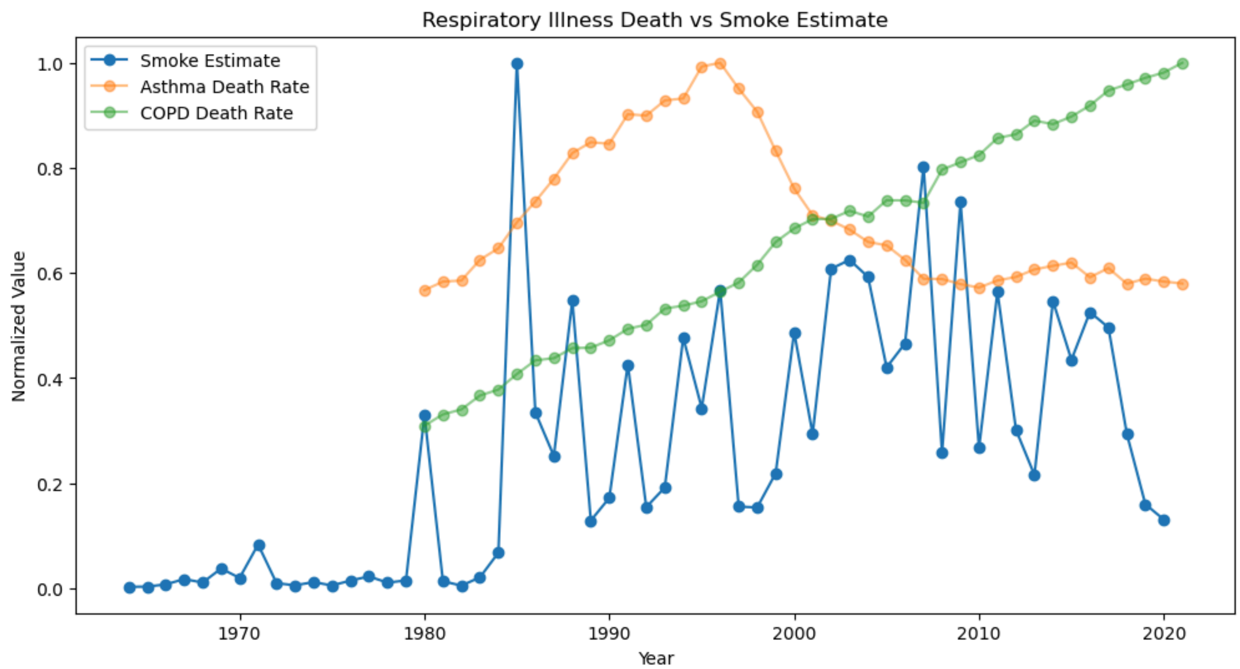

Figure 7: Trends of Respiratory Illness Death Rates vs Smoke Estimates

4.4 Forecasting Respiratory Illness Factors

The forecasts for asthma's factors are shown in Figure 8. The trend for asthma incidence seems generally reasonable. The Incidence rate forecast assumes that the massive jump the last year in the data may not affect the trend, but just results in a new starting point, which may not be prudent, but seems reasonable. However, the trend of asthma death rate is quite surprising. The forecasting model states that the death rates are supposed to increase in the next few years, with the rate of increase slightly lowering in the last few years of the forecast.

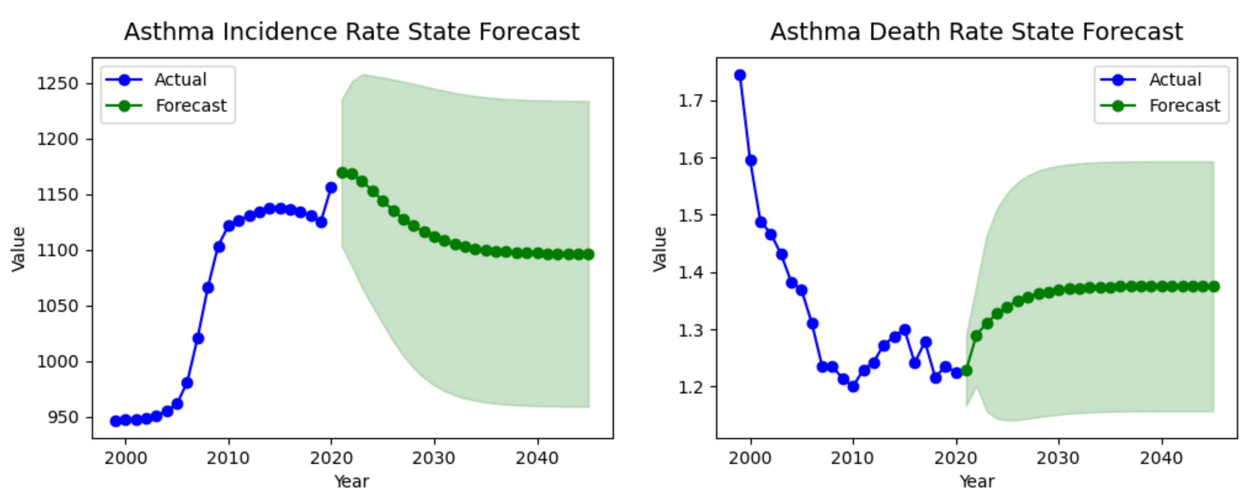

Figure 8: Forecasted rates of Incidence and Death for Asthma

The forecasts for COPD's factors are shown in Figure 9. These forecasted trends of COPD seem more reasonable considering the forecasted smoke estimates. While the COPD rates continue to increase steadily, their rates of increase are lower, perhaps due to the fact that the smoke estimates are predicted to be nearly constant by the smoke estimate forecasting model.

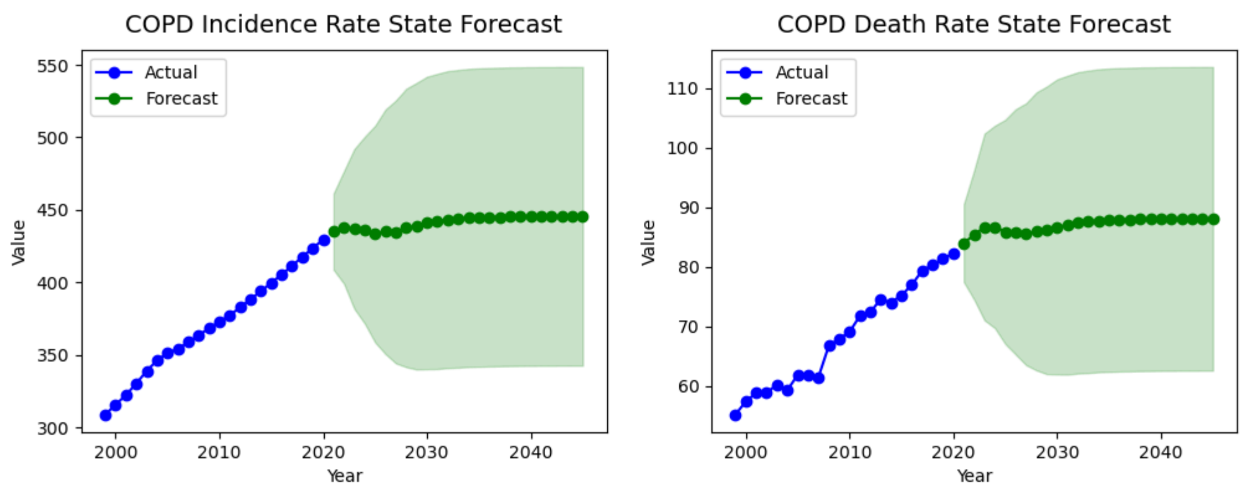

Figure 9: Forecasted rates of Incidence and Death for COPD

All factors, with the exception of the asthma incidence rate, are predicted to rise in the coming years, according to the VARMAX forecasts of each respiratory health indicator. Remarkably, the projection exhibits minimal deviation from the anticipated seasonality of the data.

5. Discussions and Implications

5.1 Implications

This analysis highlights critical insights into the interplay between wildfire smoke and respiratory health in Indianapolis, Indiana. Despite the rising trends in wildfire smoke estimates over recent decades, the overall correlation with respiratory illness rates such as asthma and COPD remains weak. However, the forecasted increase in asthma-related deaths, albeit modest, presents a potential concern, particularly for vulnerable populations.

The findings indicate a dual need for vigilance and action. While overall respiratory health risks do not appear to be escalating dramatically, the potential rise in asthma-related deaths warrants specific attention. Policymakers and public health officials should focus on:

1. Protecting Vulnerable Populations: Individuals with pre-existing respiratory conditions, children, and the elderly are most at risk. Public health initiatives must prioritize these groups for education, resources, and interventions.

2. Strengthening Air Quality Measures: Continued investment in monitoring and controlling air pollutants, particularly those directly linked to wildfire smoke, is essential to sustain recent gains in air quality.
3. Improving Health Data Collection: More granular and locally specific data on respiratory health and wildfire smoke exposure are needed to better understand and address the underlying health impacts. Currently, we are forced to focus on state-level data for our analysis, however, in the future, as Marion County ramps up its data collection efforts, we would be able to afford more nuanced analysis.

The anticipated rise in asthma-related deaths underscores the importance of sustained attention to respiratory health in Indianapolis. While immediate large-scale interventions may not be necessary, targeted measures should be implemented now to prevent more severe outcomes in the future. A phased approach, beginning with cost-effective solutions and scaled up as needed, will allow policymakers to adapt based on emerging data and trends.

By proactively addressing these issues, Indianapolis can safeguard the health of its residents, particularly those most at risk, while continuing to make progress in managing its air quality and mitigating the broader impacts of wildfire smoke. The findings reinforce the need for ongoing vigilance and collaboration to ensure the city's preparedness for evolving environmental challenges.

5.2 Reflection

Every decision that I made in this project was guided by human-centered data science, ensuring that my analysis prioritized the well-being, needs, and accessibility for Indianapolis residents. I actively applied principles of fairness, accountability, transparency, and explainability to create a project that is both ethical and actionable.

I held myself accountable by using openly available, reliable datasets from sources like the EPA and USGS. I made sure to reach out to MCPHD to get permission to access their data since they did not have a license mentioned on their site.

I carefully documented every step of my analysis, from preprocessing to model selection, to ensure anyone could verify or replicate my work. When choosing forecasting models like ARIMA and VARMAX, I selected them because they are well-established and their mechanisms are easy to explain. I prioritized explainability by designing figures and writing about my findings in a way that people without technical expertise could understand. For example, I used clear visualizations to show trends in respiratory health and wildfire smoke, avoiding unnecessary jargon. When presenting forecasts, I made sure to explain what the results mean in practical terms for the community.

Through this project, I aimed to do more than just analyze data. I wanted to empower Indianapolis residents and decision-makers with actionable insights. By actively applying human-centered principles, I made sure that my work respected the community, and provided recommendations to help protect public health in a meaningful and equitable way.

6. Limitations

The analysis of this project faces several significant limitations due to the inherent complexity of the task, the availability and quality of data, and the constraints of time and resources. These limitations impact both the scope and depth of the findings and should be carefully considered when interpreting the results:

6.1 Data Issues

The use of historical data introduces challenges related to the consistency and reliability of records. Variations in reporting practices, changes in data collection methodologies, and technological advancements over time can create inconsistencies. For example, the IHME dataset provides state-level data that lacks local granularity, limiting the ability to make precise conclusions about health impacts specific to Indianapolis. Furthermore, the Marion County data available from the MCPHD dashboard covers only a limited time span (2016–2023), making it infeasible for usage, thus leaving gaps in long-term trend analysis.

6.2 Smoke Estimate Calculation Issues

The smoke estimate model relies on fire size, type, and distance, ignoring critical factors such as wind patterns, atmospheric pressure, and fire intensity. These unaccounted-for variables significantly influence smoke dispersion and its health impacts. Additionally, the assumption that prescribed burns contribute less smoke compared to wildfires is generalized and does not consider variations in burn conditions. The absence of data for international wildfires (in the southern Canadian regions), which could affect smoke levels in Indianapolis, further limits the accuracy of the smoke estimates.

Moreover, the use of AQI data as a validation metric for smoke estimates is inherently flawed. AQI reflects overall air quality and includes factors unrelated to wildfire smoke, such as industrial emissions and vehicle pollution. As such, while AQI provides a rough proxy for smoke-related air quality, it cannot serve as a definitive benchmark for validating smoke estimates.

6.3 Forecasting Model Limitations

The forecasting models used, including ARIMA and VARMAX, are limited by small sample sizes and inherent assumptions. For instance, the dataset for respiratory health indicators spans just 20 years (1999–2020), making forecasts highly sensitive to outliers. The low-resolution annual data further restricts the ability to capture short-term anomalies caused by individual wildfires. Additionally, the stationarity assumption of these models may not hold, rendering the results less reliable. Moreover, it is hard to assess the impact of major events on respiratory illnesses such as COVID-19 since the dataset only spans until 2020, not covering the entire timespan of the pandemic.

6.4 Correlation vs Causation

This analysis is observational and limited to identifying potential associations between smoke estimates and health outcomes. While significant correlations may suggest trends, causation cannot be inferred. The uncertainties in estimating both the smoke exposure and health metrics undermine the validity of the effect sizes observed in the analysis.

Despite all these limitations, the analysis provides a foundation for understanding the relationship between wildfire smoke and respiratory health in Indianapolis. Future work addressing these limitations could yield more accurate and actionable insights.

7. Conclusion

This analysis explored the health impacts of wildfire smoke on the residents of Indianapolis, Indiana, with a particular focus on asthma and COPD-related health outcomes. By leveraging data from various sources, including wildfire occurrences, air quality indices, and state-level respiratory health statistics, the study sought to understand the trends and potential future implications of smoke exposure on public health. Using human-centered data science principles, the research aimed to inform policy and intervention strategies for the city's public health response.

Key findings reveal a complex relationship between wildfire smoke estimates and respiratory health indicators. While the occurrence of wildfires and their associated smoke levels have increased over the past decades, correlations with state-level respiratory health metrics such as incidence and mortality rates were found to be weak. This suggests that other factors, including improved air quality control measures and advances in medical care, may be mitigating some of the potential health impacts in Indianapolis. Forecasting models predict a relatively stable trend in smoke estimates over the next 25 years, while certain respiratory outcomes, like asthma-related deaths, are projected to rise in the next 10 years, highlighting a need for targeted attention.

The findings of this study have important implications for public health and policymaking. Policymakers should prioritize the protection of vulnerable populations, such as children, the elderly, and individuals with pre-existing respiratory conditions, by implementing targeted interventions and increasing community awareness during wildfire seasons. Additionally, sustained investments in air quality monitoring and proactive health data collection at the local level can enhance Indianapolis's resilience against future environmental challenges.

By applying a human-centered approach, this project demonstrates the value of integrating diverse data sources and analytical methods to address pressing public health issues. While the results serve as a foundational understanding of wildfire impacts in Indianapolis, they also highlight the need for continued research, collaboration, and action to safeguard the health and well-being of the city's residents.

8. References

1. *Indianapolis*. (n.d.). Wikipedia. Retrieved 11 28, 2024, from https://en.wikipedia.org/wiki/Indianapolis
2. *Why Wildfire Smoke is a Health Concern*. (n.d.). US EPA. Retrieved 11 28, 2024, from https://www.epa.gov/wildfire-smoke-course/why-wildfire-smoke-health-concern
3. *Health Effects Attributed to Wildfire Smoke | US EPA*. (2024, October 23). Environmental Protection Agency (EPA). Retrieved November 28, 2024, from https://www.epa.gov/wildfire-smoke-course/health-effects-attributed-wildfire-smoke
4. *How Wildfires Affect Our Health*. (n.d.). American Lung Association. Retrieved 11 28, 2024, from https://www.lung.org/blog/how-wildfires-affect-health
5. *Particle Pollution*. (2023, October 25). American Lung Association. Retrieved November 28, 2024, from https://www.lung.org/clean-air/outdoors/what-makes-air-unhealthy/particle-pollution
6. 2021 Fine Particles (PM2.5) Data Summary Report. (n.d.). *Indiana Department of Environmental Management*. https://www.in.gov/idem/airmonitoring/files/monitoring_summary_pm_2021.pdf
7. The relationship of respiratory and cardiovascular hospital admissions to the southern California wildfires of 2003. (n.d.). *National Library of Medicine*. https://doi.org/10.1136/oem.2008.041376
8. *Combined wildland fire datasets for the United States and certain territories, 1800s-Present (combined wildland fire polygons)*. (n.d.). USGS. https://www.sciencebase.gov/catalog/item/61aa537dd34eb622f699df81
9. *AQS API | AirData | US EPA*. (2020, January 10). U.S. EPA Web Server. Retrieved November 28, 2024, from https://aqs.epa.gov/aqsweb/documents/data_api.html
10. *All people living long lives in full health*. (n.d.). IHME. https://www.healthdata.org
11. *Data Use Agreement*. (n.d.). IHME. https://www.healthdata.org/Data-tools-practices/data-practices/ihme-free-charge-non-commercial-user-agreement
12. *Autoregressive integrated moving average*. (n.d.). Wikipedia. https://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average

9. Data Sources

1. *Combined wildland fire datasets for the United States and certain territories, 1800s-Present (combined wildland fire polygons)*. (n.d.). USGS. https://www.sciencebase.gov/catalog/item/61aa537dd34eb622f699df81
2. *AQS API | AirData | US EPA*. (2020, January 10). U.S. EPA Web Server. Retrieved November 28, 2024, from https://aqs.epa.gov/aqsweb/documents/data_api.html
3. *VizHub - GBD Results*. (n.d.). Interactive data visuals. Retrieved November 28, 2024, from https://vizhub.healthdata.org/gbd-results/
4. Zuker, J. (n.d.). *Explore Data*. Marion County Public Health Department. Retrieved November 28, 2024, from https://marionhealth.org/explore-data/
5. *Federal Reserve Economic Data*. (n.d.). https://fred.stlouisfed.org