DATA 512 – Project Report
Fang Yu Lim

Health Impact of Wildfire Smoke on Wichita, Kansas

Introduction

Kansas faces an increasing threat from wildfire smoke, a growing environmental hazard exacerbated by climate change. This analysis focuses on the impact of wildfire smoke on Wichita, a key urban and economic hub, and explores the broader implications of persistent exposure to wildfire smoke on the state's public health. Wildfire smoke is not confined to areas near active fires; it travels across great distances, degrading air quality far from its origin. By understanding how continuous exposure to wildfire smoke affects health and the economy, Wichita can better plan prevention strategies and mitigate future risks.

Kansas's economy relies heavily on agriculture and aircraft manufacturing, industries particularly vulnerable to degraded air quality. Wildfire smoke poses serious risks to outdoor work in agriculture, increasing the likelihood of respiratory issues, eye irritation, and other health problems. These health concerns reduce workforce safety and productivity and can potentially impair crop growth due to sunlight reduction, which can lower yields and create financial hardships for farmers. Similarly, the manufacturing sector faces challenges when contaminants from wildfire smoke infiltrate manufacturing facilities, contaminating equipment and products and leading to higher defect rates. Severe air quality may even force temporary shutdowns, resulting in lost revenue and reduced productivity.

The health impacts of persistent wildfire smoke extend beyond the workforce, affecting entire communities. Airborne particles can exacerbate preexisting conditions such as asthma and health diseases. Prolonged exposure is associated with health problems such as the increased risk of respiratory infections, cardiovascular events, and long-term lung diseases. Studies also suggest that sustained exposure can lead to the development of asthma in children. These health challenges increase demand for healthcare services, potentially straining local resources.

While existing research has broadly examined the health impacts of wildfire smoke, few studies focus specifically on Kansas. This analysis seeks to address that gap, offering actionable insights to inform local policymakers and city councils on strategies to mitigate the risks of wildfire smoke. Wichita, known as the "Air Capital of the World," is an ideal case study due to its economic significance, diverse population, and geographic vulnerability.

With a population of 397,252 (2020 Census) and a metropolitan area exceeding 647,000, Wichita exemplifies the interconnectedness of urban and rural communities in Kansas. Sedgwick County, home to Wichita, is among the state's most densely populated areas, making its residents particularly susceptible to the health and economic impacts of wildfire smoke.

As Wichita's population continues to grow, understanding the potential effects of wildfire smoke is critical to safeguarding public health, and preparing for escalating challenges. This research matters because it provides a foundation for more effective prevention and mitigation strategies tailored to Kansas's unique needs.

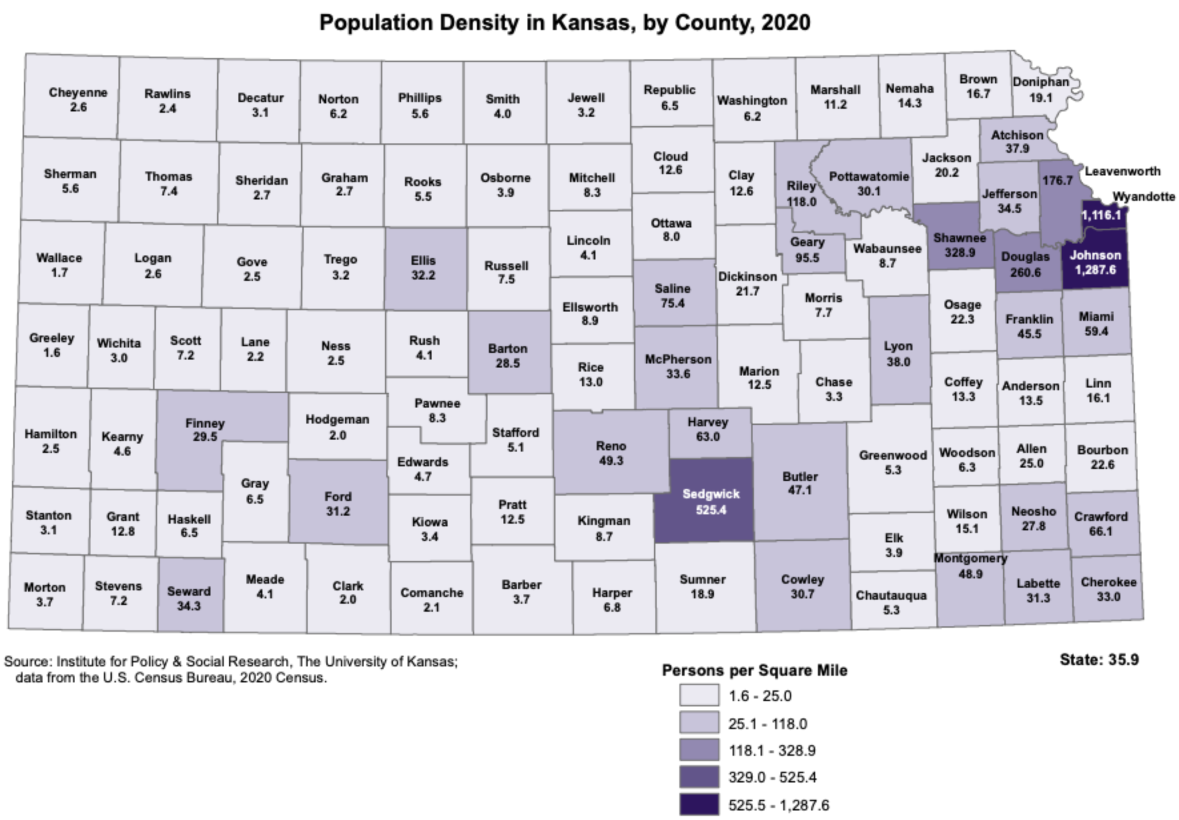

Figure 1: A map displaying the population density in Kansas, by county in 2020.

Background/Related Work

Background

According to the Environmental Protection Agency (EPA), the main pollutant emitted from wildfire smoke is fine particles, also known as $PM_{2.5}$. These are particles with a diameter of

2.5 or smaller and are of great concern to public health as they can travel deep into the lungs and may even enter the bloodstream. Existing studies from the CDC and EPA provide a broad understanding of how exposure to wildfire smoke contributes to respiratory illness and other long-term health risks. However, there is a gap in research specific to Kansas. This analysis aims to fill that gap by focusing on localized impacts. These studies helped narrow down health concerns to investigate.

Research Questions

There are 3 main questions that guided this analysis:

1. What are the estimated impacts of wildfire smoke on Wichita?
   This question seeks to assess changes in smoke levels over time, to quantify the extent of exposure in this region.
2. Is there a relationship between wildfire smoke levels and asthma-related emergency department visits?
   Given that there has been research that shows that wildfire smoke worsens respiratory conditions, this question examines whether increased exposure to wildfire smoke correlates with a rise in asthma-related hospital visits in Wichita.
3. Are there correlations between smoke levels and respiratory organ cancer or lung cancer rates?
   This question explores the long-term health effects of persistent exposure to wildfire smoke, focusing on the development of cancer in the respiratory system. The goal is to understand if smoke exposure is linked to increased rates of these conditions in Wichita.

My initial hypothesis was:

1. There is an increase in wildfire smoke in Wichita, and the trend is growing.
2. As wildfire smoke levels increase, so do the number of visits to the emergency visits due to asthma.
3. There is a moderately strong relation between smoke levels and respiratory or lung cancer rates.

Data

To estimate and predict the impact of wildfire smoke, I utilized multiple datasets sourced from trusted agencies like the U.S. Geological Survey (USGS), the U.S. Environmental Protection Agency (EPA), the U.S. Fire Administration, the Kansas Department of Health and Environment (KDHE), and the Centers for Disease Control and Prevention (CDC).

DATA 512 – Project Report
Fang Yu Lim

A more detailed description of how to access the data and variable descriptions is available in the accompanying modules and README file.

1. **Wildland Fire Data**
   To estimate the impact of wildfire smoke, I used the combined Wildland Fire Dataset from the USGS, covering 1961 to 2020. This dataset is available in GeoJSON format and includes wildfires and prescribed fires across the U.S. I used the "combined" version that removed duplicates. Key variables included geospatial polygons indicating the location of the fires, the number of acres burned, the year of occurrence, and the fire type (e.g., wildfire or prescribed fire).
2. **Air Quality Index (AQI) Data**
   Historical AQI data was retrieved by making an API call through the Air Quality Service (AQS) API by the EPA. For this analysis, I retrieved the records of gaseous AQI pollutants (CO, SO_2, NO_2, O_2) and particulate AQI pollutants (PM_{10}, $PM_{2.5}$, and acceptable $PM_{2.5}$). The data did not align perfectly with the wildland fire dataset and had missing data. As AQI is a common measure of air quality, I used it as a guide for the accuracy of my analysis.
3. **Fire Death and Injury Data**
   To explore immediate health implications such as death and injury related to wildfire, I obtained the *Fire Death and injury rates (2013 – 2022)* dataset from the U.S. Fire Administration. Although this dataset is not specific to Kansas I thought it would be interesting to study. However, further investigation revealed several limitations. The dataset only contains data for 2022, it doesn't specify whether deaths or injuries were directly related to wildfires and lacks detailed causes. Due to these issues, I decided not to draw conclusions related to wildfire smoke based on this dataset. Nonetheless, I retained this dataset in my notebook for reference, as it could provide a useful snapshot of overall fire-related death and injury trends. The details of my findings will not be discussed in this report.
4. **Asthma Data**
   Another immediate, and semi-chronic health implication I explored was asthma. I retrieved the data from the KDHE. While the data was available on a county level, it was only available from 2008 to 2018. Despite the limited data, it was of the lowest granularity to Wichita that was available. Therefore, I decided to make an analysis based on it.
5. **Cancer Data**
   The cancer data was retrieved from the CDC, it contains detailed cancer statistics and covers cancer incidence rates for various cancer sites across the United States

and Puerto Rico. I specifically filtered for only data for Kansas. The data is available from 19999 to 2021.

Model Selection

For this project, I employed Prophet, a forecasting model developed by Meta, to make predictions. Prophet was selected for its suitability for time-series forecasting and effectiveness in handling missing values and outliers. While I also explored linear regression and non-linear regression models, Prophet resulted in the best performance in capturing the trends and seasonal patterns in the data.

Methodology

Wildfire smoke

For this analysis, I focused on fires within a 650-mile radius of Wichita, according to the scope assigned. Focusing on the fires between 1961 and 2021, inclusive, we are left with 26044 observations.

Smoke is dependent on multiple factors such as acres burned, wind direction, and wind intensity. However, for this project, I only had access to the acres burned, the distance from Wichita, and the type of fire. Prescribed fires are sometimes called "controlled burns". A set of guidelines is determined to determine the environmental conditions under which it will burn. This allows us to prevent a destructive fire while creating diverse habitats for plants and animals. Therefore, it is safe to assume that prescribed fire will cause less smoke. It is also known that the larger the number of acres burned the more smoke we expect, and the further we are from a fire, the less smoke we expect.

I started off using a simple formula to estimate the smoke impact:

$$Estimated\ Smoke = \frac{acres\ burned}{distance\ from\ Wichita(miles)}$$

This approach captures the basic relationship between the size of the fire and its proximity to Wichita, though it doesn't account for other variables that may affect the actual distribution of smoke. I opted for this simple formula due to its transparency and ease of implementation. This approach allowed for reasonable approximations to quantify estimated smoke aligning with a common understanding of wildfire impacts.

Later, I used the estimated smoke values to make forecasts, further building on the smoke impact estimation for future analysis.

DATA 512 – Project Report
Fang Yu Lim

I used the air quality index to guide the accuracy of my estimated smoke impact. Since there were multiple variables included such as particulate matter and gaseous pollutants, I opted for the most straightforward way: grouping the data by year and calculating the average AQI for the year, regardless of the individual pollutant parameters. While there are other methods, such as analyzing specific pollutants or using weighted averages, this approach provided a simple and clear representation of air quality trends over time. It offered a reasonable approximation of the effects of wildfire smoke on Wichita's environment. I will keep in mind the differences in patterns between the AQI and my smoke estimates when interpreting correlations and patterns discovered. *Note: The estimated smoke was multiplied by 100 when visualizing to be on a similar scale as AQI to allow for easier comparison*.

To capture the relationships between multiple variables and forecast future trends, I built several models. Initially, I used a linear regression model, incorporating year, acres burned, distance from Wichita, and fire type. I mapped fire types to weight values based on how likely they are to produce smoke with the following weights:

Wildfire: 0.4, Likely Wildfire: 0.3, Prescribed Fire: 0.2, Unknown – Likely Wildfire: 0.15, Unknown – Likely Prescribed Fire: 0.05

These weights reflect common assumptions that wildfires produce more smoke than prescribed fires, to refine the model's predictions.

Since linear regression can sometimes lead to negative predictions, which are not meaningful in the context of smoke levels, I explored two alternative approaches to ensure all predictions were non-negative.

The first approach was to use a non-negative least squares (NNLS) model. This is a variation of linear regression that ensures that all predictions remain non-negative.

The second approach involved shifting the predictions so that any negative values were corrected. I did this by applying a shift to all predictions by the absolute of the most negative value. This ensured that predictions were all positive and maintained their relative scale. This approach allowed the model to make meaningful forecasts while preserving the structure of the data.

Given the expected impacts of global warming, I simulated a dataset where more acres are expected to burn each year. I set the base number of acres burned as the average value from the dataset and increased it by a fixed amount of 175 acres annually, adding noise to simulate variability. For fire types, I assumed an increasing frequency of wildfires, adjusting the probability of a fire being classified as a wildfire to be higher than that of a prescribed fire. This assumption does not account for potential precautions like increasing prescribed fires to prevent wildfires. For the distance, I randomly sampled distances between 0 and 650 miles from Wichita, reflecting the expected spread of smoke.

An alternative approach could involve sampling distances based on the distribution of total acres burned each year, aligning fire impact more closely with the actual distribution of fire events. This simulated dataset aimed to generate realistic forecast data, which provided a basis for evaluating the performance of the models in capturing trends and making predictions.

A third model I explored is Prophet, a time-series forecasting model developed by Meta. For this approach, I used a Linear Regression model with years as the predictor to predict future acres burned, capturing long-term trends. I also considered the shape of my distance data by modeling future distances using a log-normal distribution. This distribution is suitable for positively skewed data like distance, where values cannot be negative. By adjusting the distribution's shape and scale parameters, I simulated realistic future distances, reflecting the variability in fire locations relative to Wichita. This approach integrates long-term trends and provides a comprehensive framework for forecasting wildfire smoke impacts.

Asthma

Since I only had yearly counts of asthma-related emergency department visits, I considered two approaches for analysis. The first was to check for a correlation between the yearly counts and estimated smoke. However, this approach was not feasible due to the different time granularities of the data.

Thus, my second approach involved calculating an average smoke estimate per year and then examining its correlation with the annual counts of asthma-related emergency visits. This method allowed for a more consistent comparison between the two datasets.

To forecast the number of asthma-related emergency visits due to smoke levels, I used the Prophet model and experimented with three methods of calculating smoke. Each method aimed to capture different aspects of the relationship between smoke exposure and asthma visits.

1. Per incident: This approach treated each fire incident individually. The assumption was that the impact of each fire could be analyzed separately to see how individual fires correlate with emergency department visits. However, this method was too granular and is not supported by the asthma data we have access to.
2. Average smoke per year: This approach smooths out variability of smoke levels across incidents and helps account for overall trends allowing for a clearer view. By

averaging smoke exposure, this method is better aligned with the time granularity of asthma visit data.

3. The sum of all smoke in a year: I summed all the smoke estimates from all fires within a year. This approach captures the cumulative impact of multiple fire events over a single year. This approach may be more informative to identify trends in the overall health impact over time.

Cancer

For cancer analysis, I focused on respiratory system cancers, particularly lung cancer. I aimed to examine the correlation between estimated smoke levels and the following four cancer types:

- Nose, Nasal Cavity, and Middle Ear
- Larynx
- Lung and Bronchus
- Respiratory System

I am to explore whether prolonged exposure to wildfire smoke is associated with increased rates of these conditions. This could provide insights into the long-term health risks associated with exposure to wildfire smoke.

Findings

Wildfire patterns

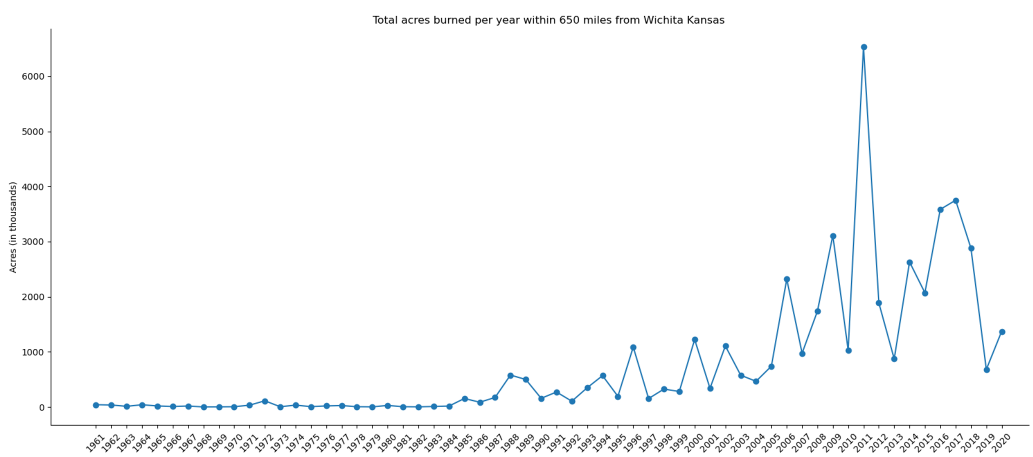

Figure 2: Total acres burned per year within 650 miles of Wichita Kansas

This time series graph shows the total acres burned per year, for wildfires up to 650 miles from Wichita Kansas. The x-axis represents the year, and the y-axis represents the total acres burned in thousands of acres. The graph indicates a significant increase in the number of acres burned per year, this trend suggests a growing risk of wildfires in the region.

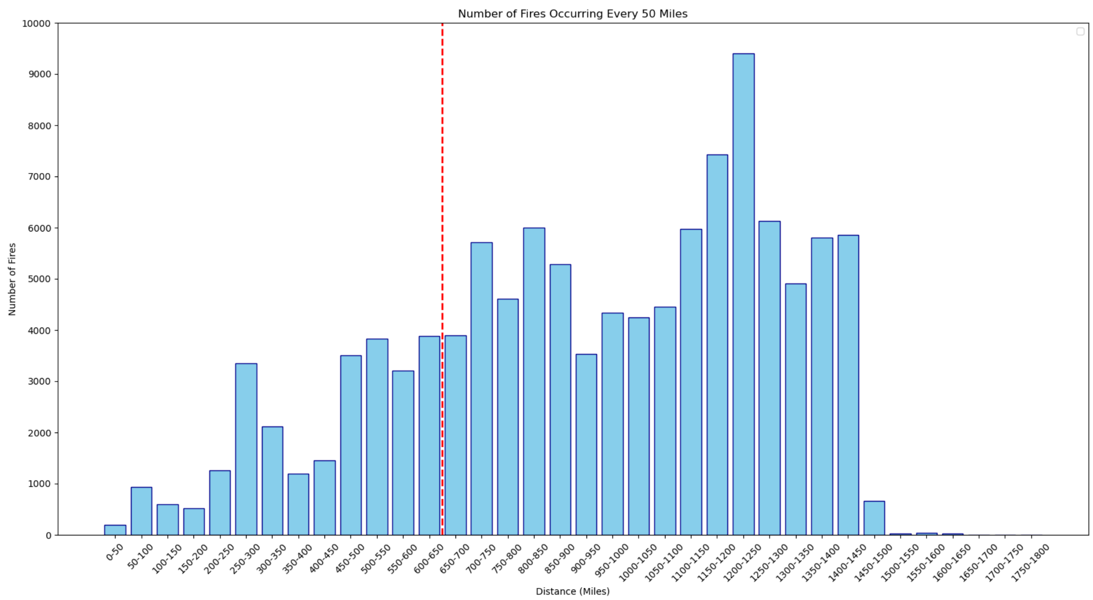

Figure 3: The number of fires occurring every 50 miles from Wichita

This histogram displays the distribution of wildfires occurring in 50-mile increments, with distances ranging up to 1800 miles from Wichita. The x-axis represents the distance bins (grouped in intervals of 50 miles), and the y-axis shows the number of wildfires within each bin. The plot reveals that most wildfires are occurring more than 650 miles away from Wichita. Since the scope of this project focuses on fires within 650 miles, we observe that the section of the plot corresponding to this range is right-skewed. This suggests that within the 650-mile radius, there are fewer fires closer to Wichita, with the frequency of fires increasing as the distance grows.

Smoke Estimates

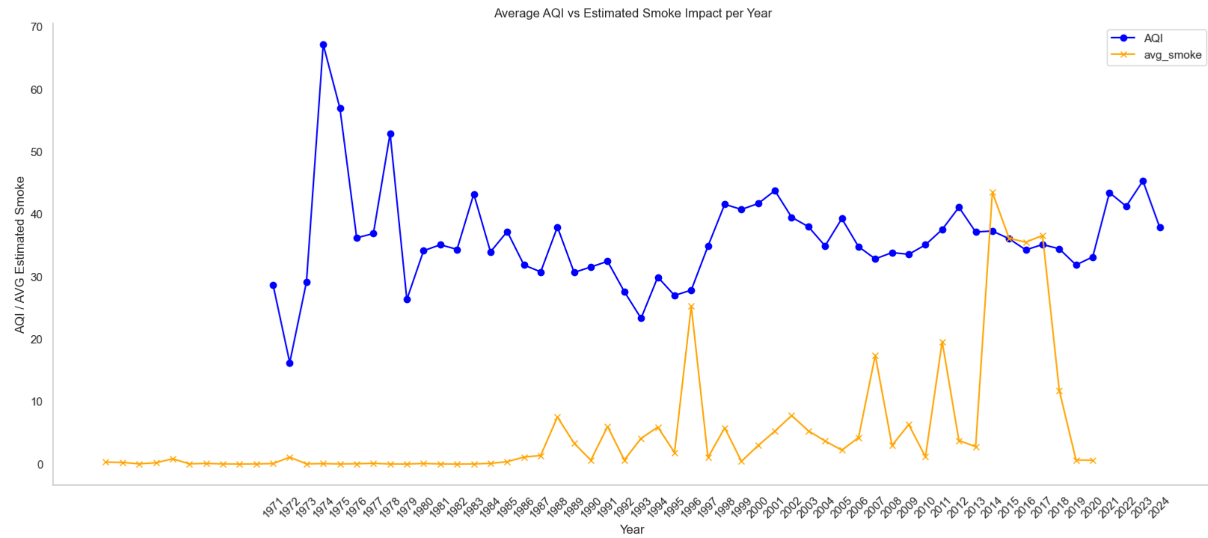

Figure 4: A time series plot between average AQI and estimated smoke impact per year.

Observing the graph between the average air quality index and estimated smoke impact per year generated with our simple formula, we only see a weak correlation. This suggests that our formula for estimated smoke might not be capturing key relationships, we might be overestimating the air quality index by using an average, or we are not including some important features. After further research, we found out that
Kansas is known to have strong winds, therefore, future studies should consider incorporating additional factors like wind speed and direction to better understand smoke impacts. Stronger winds may cause wildfire smoke to disperse faster, leading to a low observation of particulates in the air.

The Pearson correlation coefficient between the two is – 0.051 which proves that there is a negligible linear relationship between average air quality index and average smoke impacts. However, it is worth noting that particulate AQI was not observed until the late 1980s, and this could affect our AQI scores. $PM_{2.5}$ is the primary component of wildfire smoke with the most severe health implications. Without this data, AQI trends may not accurately reflect smoke-related air quality.

Linear Regression Model

```
Linear Regression Coefficients (may allow negative):
Year: -0.001004042459986597
Acres: 8.10024379883744e-05
Distance: -0.0008544412484807517
Fire_Weight: -0.8802468920427424
Linear Regression Intercept: 2.621978289649488
Mean Squared Error: 14.888490148426204
```

Our linear regression model suggests that there is a minor decrease in smoke impact over time, larger fires tend to increase smoke impact, smoke impact decreases as distance from fire increases, and that the severity of the fire has a substantial effect on smoke impact. Most of these findings align with our assumption except for the year variable. We had expected smoke to increase yearly due to global warming. However, this could be due to major spikes in the acres burnt per year.

Non-Negative Least Squares Model

```
NNLS Coefficients (Non-Negative):
Year: 0.0
Acres: 8.225207605365321e-05
Distance: 0.0
Fire_Weight: 0.0
NNLS Intercept: 0.0
NNLS Mean Squared Error: 14.977949811041668
```

This model suggests that acres burned is the only significant factor, with a positive impact on smoke impact.

Comparing our models, they have similar mean squared errors, indicating similar predictive performance. However, our linear regression models offer a more nuanced understanding of the factors influencing smoke impacts.

Linear Regression Smoke Forecast

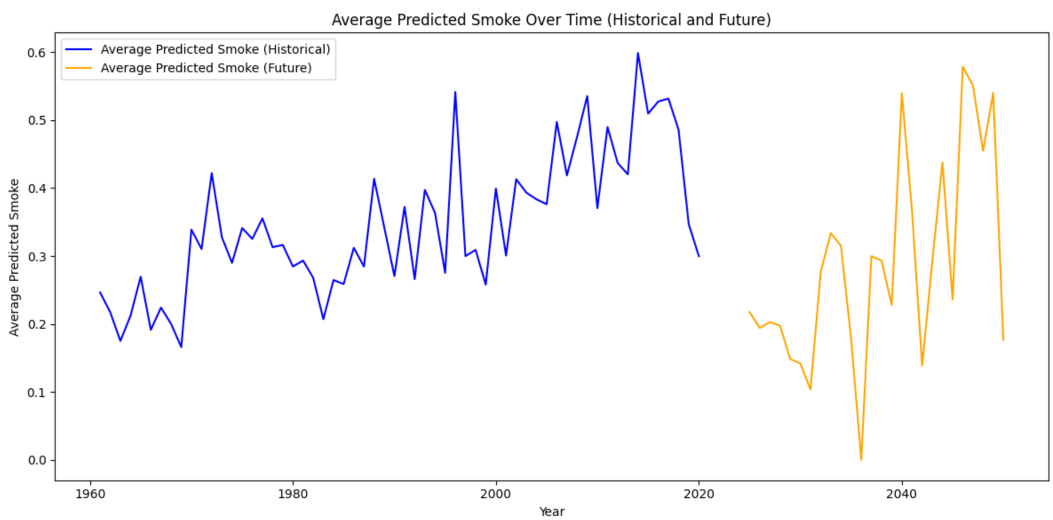

Figure 5: A time series graph showing the smoke (historical and future) predicted using the linear regression model

While the linear regression model provides a baseline for understanding historical trends, its limitations in accurately predicting future smoke levels are evident. The significant discrepancies between historical and future predictions highlight the need for more sophisticated time series forecasting techniques, such as Prophet. The average predicted smoke begins at the lowest value observed in historical data. This suggests that the way we forecasted future data for prediction was flawed.

Estimated Smoke Forecasted with prophet

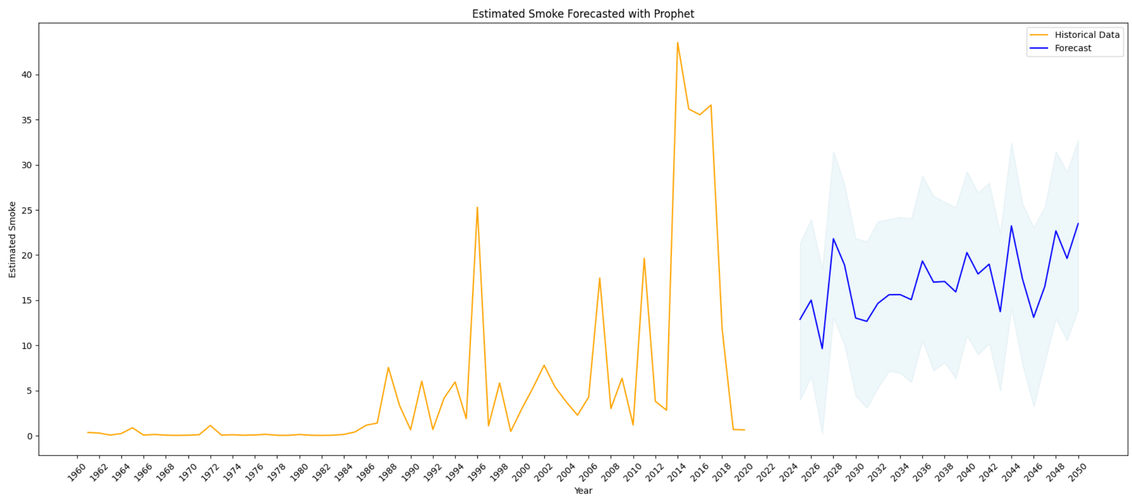

Figure 6: Time series plot of estimated smoke (using simple formula) and forecasted smoke using prophet.

The Prophet model captures a more robust forecast compared to the linear regression. The forecast suggests the smoke levels will increase over time, with a high level of uncertainty. This was expected as we had identified earlier that we should consider other factors such as wind intensity.

Health impacts

Asthma

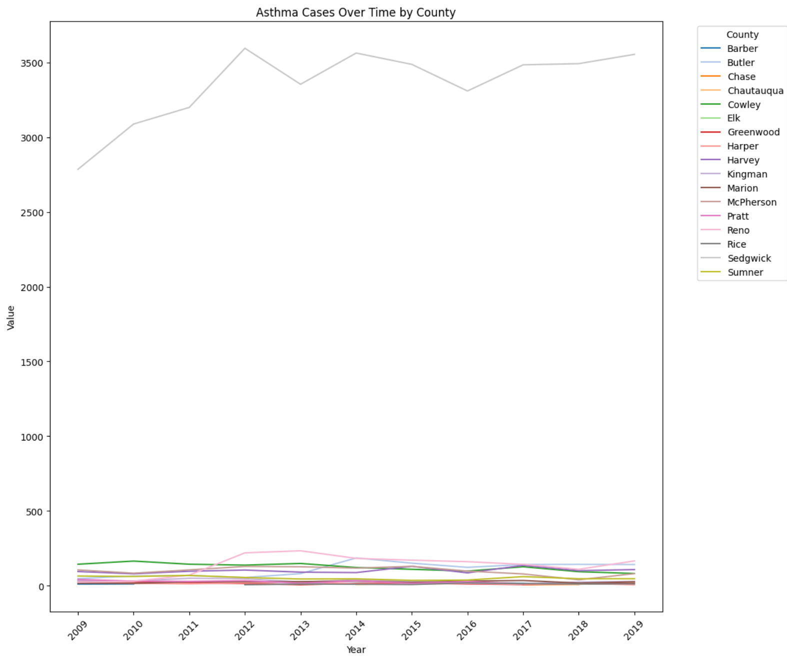

Figure 7: The number of ER visits due to asthma per year, per county.

The graph above is the number of visits to the emergency department visits for asthma for counties surrounding Sedwick County, specifically those within a 2-county radius. It reveals a concerning trend of increasing asthma cases over time in several counties surrounding Sedgwick County (when examined without Sedgwick County in the picture). This suggests a potential link to environmental factors, such as wildfire smoke exposure. It is also alarming that Sedgwick County has such a high number of emergency department visits for asthma compared to the others.

Correlation between smoke estimated smoke and ER visits (asthma)

Initial discovery showed a 0.014 correlation between each estimated smoke and the number of visits to the emergency room, suggesting a limited impact. Aggregating data can reduce noise and reveal more significant trends. Thus, I aggregated by years obtaining the average estimated smoke per year, and a stronger correlation emerged (0.299). This suggests that the number of ER visits due to asthma increases as smoke levels increase.

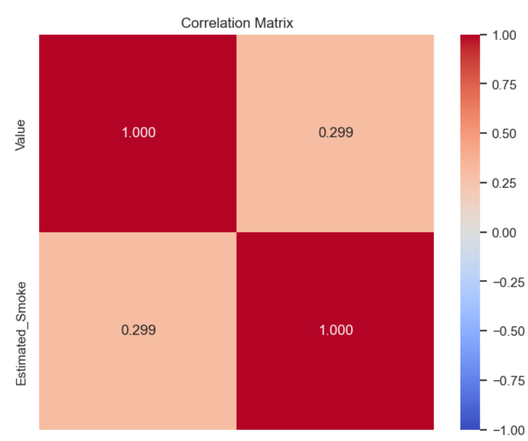

Figure 8: Correlation between estimated smoke and the number of ER visits due to asthma.

Forecasted asthma visits based on predicted smoke levels

Even though there is only very limited data available, I was interested in the accuracy of predictions that could be made about the number of ER visits due to asthma.
For this section, instead of using estimated smoke, I used predicted smoke which are the forecasted views of the Prophet model.

Approach 1: Forecasted asthma visits based on individual fire smoke emission

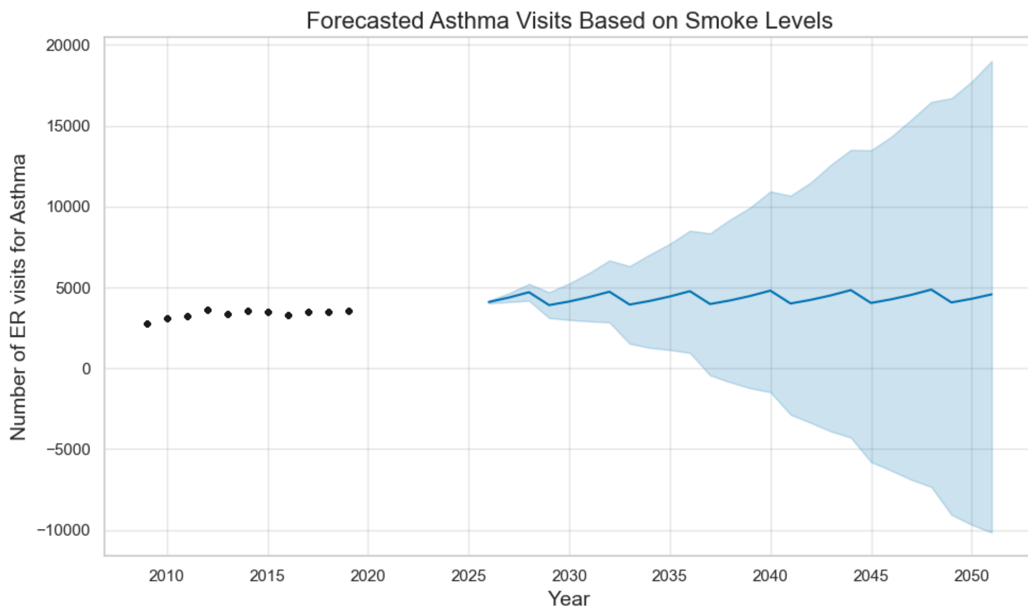

Figure 9: Forecasted asthma visits based on individual fire smoke levels

By examining each individual fire and its associated smoke emission, we observe that there is an increasing trend in asthma-related emergency room visits over the next few decades. The blue line represents the most likely scenario, while the shaded area indicates the uncertainty range. This suggests that as wildfire smoke levels increase, so will the number of people who seek medical attention related to asthma. However, since our data for asthma is a yearly aggregate, this approach would not be the most appropriate. Basing forecasts on individual fire smoke emissions might introduce too much variability that doesn't match the available health data.

Approach 2: Forecasting asthma visits based on average smoke levels

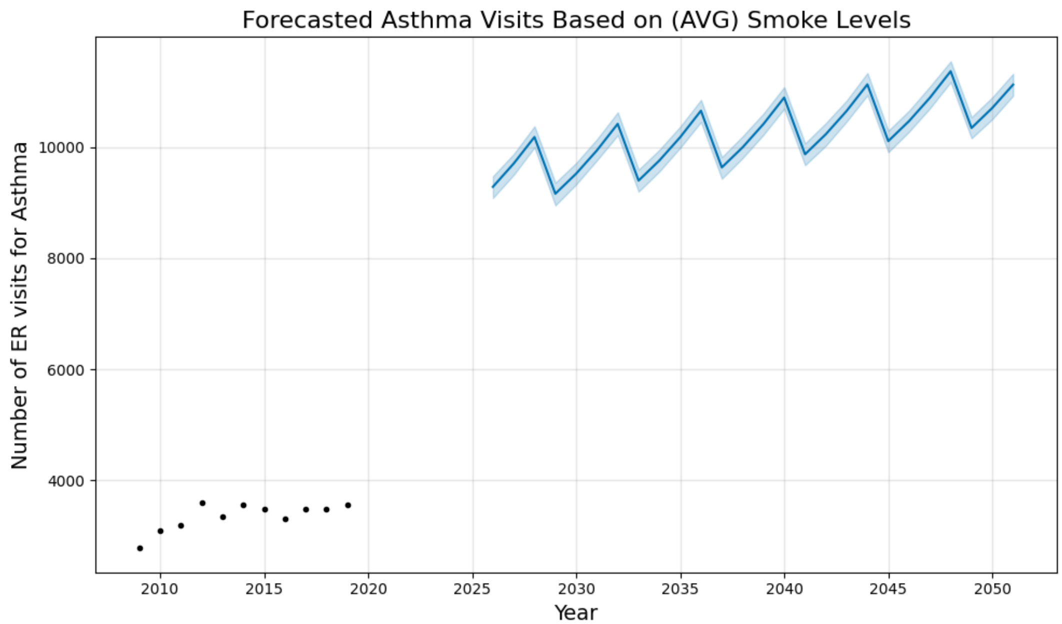

Figure 10: Forecasted asthma visits based on yearly averaged fire smoke levels

The analysis using yearly average smoke levels indicates a significant and concerning upward trend in asthma-related emergency room visits. The narrow confidence interval suggests a high level of certainty in this projection. The forecasted pattern aligns well with the observed historical trend, further emphasizing the potential for substantial increases in asthma-related health issues as wildfire smoke levels continue to rise.

Approach 3: Forecasting asthma visits based on the summed smoke levels per year

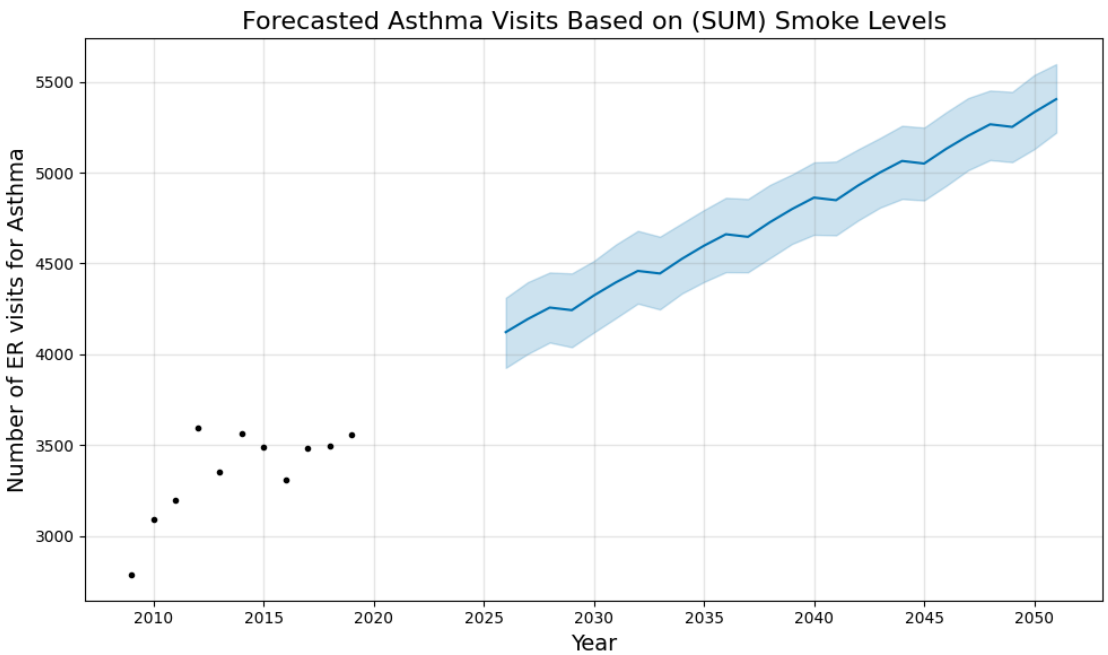

Figure 11: Forecasted asthma visits based on yearly summed fire smoke levels

The analysis using yearly summed smoke levels indicates a steady increase in the number of visits to the emergency department due to asthma. While we have a larger shaded area indicating less confidence, the overall trend suggests a concerning future for respiratory health. Despite the large uncertainty range, it reflects the relationship between wildfire smoke and ER visits because of asthma.

This approach provides a more comprehensive view of the cumulative impact of wildfire smoke on public health. However, it's important to note that the accuracy of these forecasts depends on various factors, including future climate patterns, fire management practices, and potential changes in population demographics and healthcare access.

Cancer

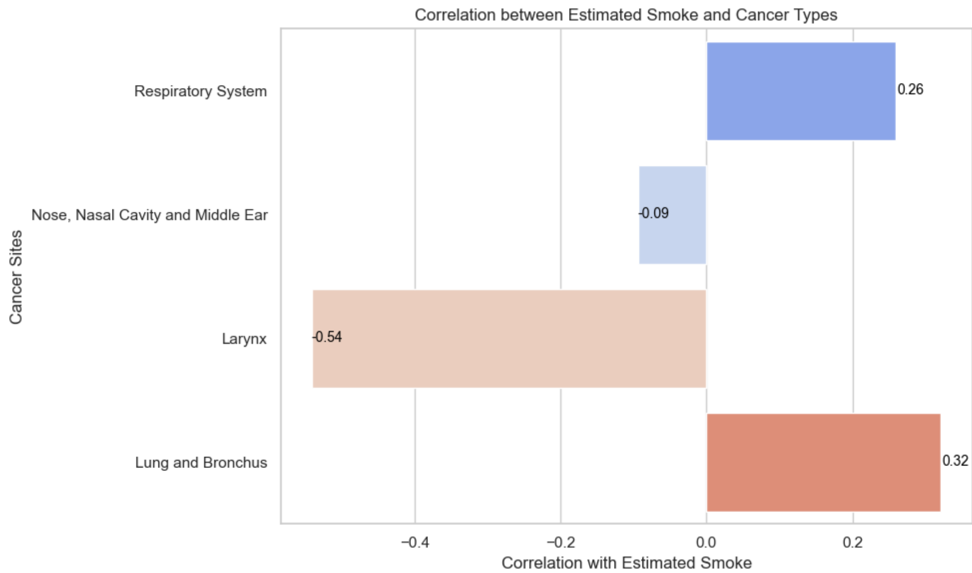

Figure 12: Correlation between estimated smoke and cancer types

The provided bar chart highlights a potential association between estimated smoke exposure and specific cancer types. Notably, there is a positive correlation between smoke exposure and the risk of respiratory system and lung/bronchus cancers. This suggests that increased exposure to wildfire smoke may elevate the risk of these cancers.

However, the negative correlation observed for larynx cancer is intriguing and requires further investigation. It's possible that other factors or confounding variables might be influencing this relationship. Cancer is also known to develop over time, a relation that is not investigated in this research. Additional research is necessary to establish a definitive causal link between smoke exposure and cancer risk, and to explore the underlying mechanisms behind these associations.

Discussion/ Implications

The findings from this project reveal a concerning trend between increasing wildfire smoke exposure and the growing health impacts in the Wichita region, particularly regarding asthma and respiratory cancers. Our analysis suggests that over time, wildfire smoke is likely to increase, potentially driven by changing climate patterns. Furthermore, the rise in smoke levels appears to correlate with an increased risk to public health, particularly in terms of a higher number of asthma-related cases. The observed increase in asthma-related emergency department visits serves as an early warning of the potential health burden that may be placed on the healthcare system in the coming decades.

The forecasted upward trend in patients needing medical attention for asthma, and correlation with respiratory system cancer underscores the urgent need for action to mitigate these health risks and prepare for future scenarios. These findings should not only warn individuals of potential health impacts but also serve as a guide for the city council to act. The city should prioritize strengthening local healthcare infrastructure and emergency response systems. Without these measures, there may be an overwhelming burden on the healthcare system during peak smoke exposure periods, resulting in individuals not receiving the care and treatments needed. Preparing for these challenges now can help ensure that the community is resilient when facing the increased wildfire smoke-related health impacts as air quality continues to worsen.

Recommendations

The need for concrete action is urgent as the impacts of wildfire smoke are forecasted to intensify in the coming decades. Therefore, based on the insights generated from this study, city officials should begin implementing changes immediately.

While the KDHE already monitors air quality for Wichita, I recommend making this data more accessible to the public, including historical data, to provide better context for current trends. The public should have access to resources to aware them of the health impacts caused by wildfire smoke, with the most useful method possibly being implementing a public outreach campaign to raise awareness about the risks.

The city should invest in the creation of a more refined model to forecast smoke impacts. This will allow for more accurate predictions and timely warnings. An alert system could warn residents of days with potentially hazardous air quality, allowing them to take protective measures in advance.

I recommend city leadership to come up with an action plan to address wildfire risks, such as promoting controlled burns to prevent wildfires and mitigate fire severity. Since many wildfires occur outside Wichita, the mayor should push for regional collaboration with neighboring cities and states on wildfire prevention.

On top of that, I recommend the city consider providing suitable masks (Eg: N-95 masks) to residents ahead of historically known periods of high smoke exposure. This would help protect the most vulnerable populations, such as children, the elderly, and individuals with respiratory conditions. Although homelessness is not a widespread issue in Kansas, the city should ensure that homeless individuals are also considered in these protective measures, as they are equally susceptible to smoke exposure.

I recommend that city leadership begin immediately and prioritize more detailed and frequent tracking of health indicators. This will allow for a better understanding of short- and long-term health impacts related to wildfire smoke. Most importantly, with the increasing emergency department visits for asthma alone, expanding access to healthcare is crucial. The city should invest in improving respiratory health infrastructure.

Things to be implemented in the short term include early warning systems, closer monitoring of healthcare concerns/ hospitalization rates, and educating the public on wildfire risks. The city should begin working on policies and collaborations targeting long-term climate change mitigation and strategies for smoke reduction. The city should begin investing in infrastructure that supports health, most likely with a focus in respiratory health due to the nature of smoke.

Reflection

Human-centered data science emphasizes the importance of considering the real-world context and the needs of people affected by the data. The goal of this project was not to analyze data but to translate findings into concrete recommendations that would help improve the health and safety of Wichita's population. This requires an understanding of the potential needs of the community to ensure that recommendations are relevant and actionable. This is related to one key principle of human-centered data science – contextual relevance – which ensures that insights derived from data align with real-world issues.

In line with the FATE principle, I made sure to not include any personally identifiable information, respecting privacy and protecting individuals.

Another key principle is transparency and interpretability. In this project, I aimed to present findings clearly and concisely and chose models and graphs which are easily interpretable or understandable. Instead of overwhelming users with code or graphs, we provide the model outcomes and share the limitations and complications to ensure full transparency which can allow feedback and continuous improvement. This ensures that all stakeholders can make informed decisions based on the best available evidence.

Limitations

A key limitation of this study was the availability and quality of data. Air quality data was inconsistent and incomplete. Our estimated smoke relied on estimates, and as we compared it to the AQI, we realized that we might be overlooking some important factors.

Other data, health data for example were available on different levels of granularity. For example, wildfire data was recorded per incident, but the lack of exact dates hindered the ability to retrieve wind data to enhance our model. Health data was aggregated on a yearly basis, with some available on a state level, and others available on a county level. Unfortunately, none of the data was available on a city level. This limits our ability to fully capture localized variations and health impacts. Our asthma dataset was available on a county level, but it had very limited years of data which may not be enough to provide a comprehensive analysis of the relationship between asthma and smoke exposure.

Our models – linear regression and Prophet, oversimplify the problem due to limited factors available. The linear regression model assumes that the relationship between variables is linear, and might not capture the complex, most likely non-linear relationships. Similarly, Prophet assumes the stability of historical trends, which may not hold true as interventions are carried out.

While we found correlations between smoke exposure and health outcomes, it is important to note that correlation does not equal causation. There may be other factors such as living environment, and pre-existing health conditions shaping these outcomes.

Additionally, the results may not be generalizable to all regions of Kansas with different environmental and lifestyle factors. Furthermore, this study does not account for potential future interventions which is something that may be in progress, there is no action to account for that. Therefore, this limits the ability to forecast long-term health outcomes with clarity.

Conclusion

This study aimed to investigate the relationship between wildfire exposure and health outcomes, specifically asthma and respiratory cancers in Wichita. Our research questions were: What are the estimated impacts of wildfire smoke on Wichita? Is there a relationship between wildfire smoke levels and asthma-related emergency department visits? Are there correlations between smoke levels and respiratory organ cancer or lung cancer rates?

The findings of this study reveal that as wildfire smoke exposure increases, so does the number of individuals who require medical attention for asthma. This suggests higher smoke levels may exacerbate respiratory conditions. The correlations between wildfire smoke and different types of cancer were less pronounced, but there are indications that exposure to wildfire smoke is likely to increase cancer risks.

This study highlights the importance of using data to understand real-world issues, generate insights, and provide actionable recommendations. Following human-centered principles ensures that results are relevant and actionable and protects the privacy of the community.

References

i. "Wichita, Kansas." *Wikipedia*, Wikimedia Foundation, 30 Nov. 2024, "Wichita, Kansas." *Wikipedia*, Wikimedia Foundation, 30 Nov. 2024, en.wikipedia.org/wiki/Wichita,_Kansas#:~:text=Wichita%20(%2F%CB%88w%C9%AAt%CA%83,population%20of%20647%2C610%20in%202020.

ii. University of Kansas. (n.d.). *Population density in Kansas* [PDF]. Kansas Data Access and Support Center. https://ksdata.ku.edu/ksdata/ksah/population/popden.pdf

iii. U.S. Environmental Protection Agency. (n.d.). *Health effects attributed to wildfire smoke*. U.S. Environmental Protection Agency. https://www.epa.gov/wildfire-smoke-course/health-effects-attributed-wildfire-smoke

iv. U.S. Environmental Protection Agency. (n.d.). *Air pollution and cardiovascular disease: The basics*. U.S. Environmental Protection Agency. https://www.epa.gov/air-research/air-pollution-and-cardiovascular-disease-basics#:~:text=The%20evidence%20is%20particularly%20strong,the%20chest%2C%20neck%20or%20shoulder

v. U.S. Environmental Protection Agency. (n.d.). *Why wildfire smoke is a health concern*. U.S. Environmental Protection Agency. https://www.epa.gov/wildfire-smoke-course/why-wildfire-smoke-health-concern#:~:text=Figure%201.,may%20even%20enter%20the%20bloodstream

vi. National Park Service. (n.d.). *What is a prescribed fire?* U.S. Department of the Interior. https://www.nps.gov/articles/what-is-a-prescribed-fire.htm

vii. World Population Review. (2024, December 4). Windiest States 2024. Retrieved from https://worldpopulationreview.com/state-rankings/windiest-states

Grammarly:

I used Grammarly to review my report to check for spelling and grammar errors.

ChatGPT:

I used ChatGPT to help with sentence structure and flow, making sure I wasn't jumping between ideas.

Data Sources

- Wildland Fire Dataset
- EPA AQI API
- CDC Cancer Statistics
- KDHE Asthma Dataset
- US Fire Death Dataset
- State Size Dataset