THE IMPACT OF WILDFIRES ON YAKIMA COUNTY

DATA 512 - COURSE PROJECT

RASHMIKA REDDY VOOKANTI

INTRODUCTION

The recent escalation in the frequency and severity of wildfires across the western United States has underscored the critical need for in-depth analysis and proactive response strategies, particularly in regions like Yakima, Washington. These wildfires, fueled by factors such as climate change and land management practices, emit significant smoke, adversely affecting public health, environmental quality, and the socio-economic landscape. This study, "Understanding the Impact of Wildfires on Yakima, Washington: A Comprehensive Analysis," aims to explore the extensive implications of these fires on the region.

Yakima County, a region celebrated for its diverse natural landscapes and agricultural prowess, has increasingly found itself at the frontline of these wildfire events. The pervasive smoke from these fires poses serious threats not only to the health and safety of the residents but also to the agricultural output, directly impacting the economic stability and livelihoods in the area. The repercussions of these wildfires extend far beyond immediate physical damages, influencing policy-making, urban planning, and healthcare resource allocation.

This study seeks to comprehend the full spectrum of effects that wildfire smoke has on Yakima County. By examining the trends in air quality deterioration, health impacts, and socio-economic disruptions, the research aims to provide a holistic understanding of the situation. Recognizing the varying health risks wildfire smoke presents to different demographics is crucial, as it helps in formulating targeted strategies for mitigation and management.

The insights gained from this analysis are not confined to benefiting Yakima's population alone. They serve as critical input for governmental bodies and civic institutions, guiding them towards implementing effective social and economic reforms. This research is pivotal in influencing the decisions of policymakers and community leaders, thereby

playing a crucial role in shaping future strategies to combat the adverse effects of wildfires.

In summary, this project aspires to translate comprehensive data into effective and forward-looking solutions, aiming to protect the community of Yakima from the escalating threat of wildfires. Through this analysis, we intend to empower local authorities and residents with the knowledge and tools necessary to build a more resilient and informed community.

BACKGROUND/RELATED WORK

While my research on the public health impacts of wildfire smoke is not specifically centered on Yakima County, it draws from a spectrum of studies that scrutinize the repercussions of wildfire smoke on respiratory well-being and healthcare facility admissions. Such investigations, although geographically diverse, offer crucial insights applicable to the respiratory health challenges within Yakima County. A pivotal Southern California study, for instance, has uncovered that particulate matter from wildfires, known as PM 2.5, poses a greater risk to health than similar particulates from other sources, leading to a marked uptick in respiratory-related hospital admissions.

Emerging research indicates an upward trend in wildfire occurrences due to climatic variations, which has been linked to increased smoke exposure and exacerbated respiratory conditions such as asthma, which is particularly relevant to the health landscape of Yakima County's population.

The research inquiries guiding my analysis are twofold:

1. Firstly, I am examining the accuracy of my smoke exposure estimates against actual air quality indices to determine if a significant correlation exists.
2. Secondly, given the established severe health impacts of wildfire smoke, I aim to correlate smoke exposure estimates with the incidence of asthma rates among the Yakima’s adolescent population.

METHODOLOGY

The project was methodically segmented into clear steps for ease of understanding and execution.

Step 1: Wildfire Data Sourcing

The foundation of the analysis was the extensive 'USGS_Wildland_Fire_Combined_Dataset.json' from the US Geological Survey. This massive dataset of 2.8GB was meticulously parsed using GeoJSON reader libraries, with a focus on fires post-1963 within a 1250-mile radius of Yakima, WA. To calculate smoke estimate, a nuanced formula factored in the fire's distance, the area burned, and its intensity. Different fire types were weighted according to their intensity, assigning higher values to wildfires compared to prescribed burns, reflective of their potential for smoke production.

$$SmokeEstimate = (AreaBurnt*FireIntensity)/DistanceFromCity$$

Why This Methodology:

- Objective-Oriented: Focused on extracting relevant data to understand the spatial distribution and characteristics of wildfires around Yakima.
- Comprehensive Data Scope: Chosen to capture a wide array of fires, ensuring a thorough analysis of potential impacts on air quality.
- Ethical Consideration: Emphasized recent and relevant data, aligning with ethical data use principles.

Step 2: Air Quality Index (AQI) Retrieval

Air quality data was pivotal to this study. By accessing the US AQS database through a generated API key, historical AQI data was systematically fetched. The lack of proximal weather stations necessitated the creation of a spatial bounding box extending up to 250 miles from Yakima, encompassing data from multiple stations. The yearly data collection involved averaging out AQI values to construct a comprehensive record from 1985 to 2022, despite variations in station operational years and data availability.

Why This Methodology:

- Data Integrity: Sourcing from a reliable database ensured accuracy and standardization in air quality measurement.
- Spatial Inclusivity: Expanding the data collection area allowed for a more holistic view of air quality impacts in the broader Yakima region.
- Temporal Relevance: Focusing on a substantial historical range provided insights into long-term air quality trends.

Step 3: Impact Analysis of Wildfire Data

The collected smoke estimate data was visualized to ascertain the spread and impact of wildfires in the vicinity of Yakima. Histograms illustrated the frequency of fires by proximity, and time series graphs elucidated the annual burn area trends. A critical component was the comparative analysis of the calculated smoke estimates against the AQI, revealing the air quality's fluctuation over time and the apparent influence of wildfire smoke.

Why This Methodology:

- Data Visualization: Empowered clear communication of complex data, making findings accessible to various stakeholders.
- Insightful Comparisons: Enabled the assessment of how wildfire smoke correlates with changes in air quality.
- Human-Centered Approach: Focused on translating environmental data into public health implications, emphasizing the project's relevance to the community.

Step 4: Project Extension for Public Health Insights

Building upon the gathered data, the extension of the project zeroes in on the health impacts of smoke on local adolescents. The Children's Alliance Provider dataset on asthma prevalence among tenth graders forms the crux of this analysis, offering a longitudinal view of health trends against the backdrop of environmental changes. This step interweaves health data with environmental factors, aiming to deliver a comprehensive narrative on public health preparedness and adaptive strategies in the face of evolving wildfire patterns.

Data Used:

Asthma prevalence data

Column Name	**Description**
LocationType	Type of geographical location
Location	Specific location within Washington

TimeFrame	Year
DataFormat	Here the format is percentage
Data	Prevalence percentage of Asthma in 10th graders

Why This Methodology:

- Focused Analysis: Targeting a specific demographic (adolescents) ensures a detailed understanding of how wildfire smoke impacts vulnerable groups in the community.
- Longitudinal Data: Using a dataset that spans several years allows for the observation of trends and changes over time, crucial for understanding evolving health impacts.
- Ethical Approach: By focusing on a demographic that is often more vulnerable to environmental changes, the study aligns with ethical principles of protecting and prioritizing at-risk populations.

Step 5: Analysis of asthma rates with Smoke

● In the further analysis, read the asthma data downloaded from the website and read in the notebook from the percentage_of_tenthgraders_with_asthama.xlsx and merged it with the previously calculated Smoke Estimate Data.

● Renamed 'es_aqi' to 'smoke_estimate' for clarity in the heatmap, Perform the correlation analysis using Pearson's correlation coefficient between Smoke Estimate and asthma_rate.

● Asthma Rate Trends Over Time with Smoke Estimate: This analysis compared historical and predicted AQI values, reflecting on past trends and future expectations in air quality and its potential impact on asthma rates.

Why This Methodology:

- Comprehensive Assessment: Integrating environmental and health data provides a holistic view of the wildfire smoke's impact on public health.

- Statistical Rigor: Employing correlation analysis and regression models lends scientific robustness to the study, allowing for more reliable interpretations of the data.
- Predictive Insight: The use of predictive models to forecast future trends helps in proactive public health planning and environmental management.
- Human-Centered Approach: The analysis directly addresses the health concerns of the community, particularly the youth, emphasizing the human impact of environmental changes.

Human-Centered Considerations

- Community Well-being: The methodology is centered around the health and safety of the community, particularly the younger population, highlighting the human aspect of environmental changes.
- Transparent and Accessible Reporting: The use of visual tools like graphs and heatmaps ensures that the findings are comprehensible and accessible to a broad audience, including non-experts.

FINDINGS

1. Hypothesis 1: Wildfire and Smoke Impact
    a. **Histogram of Wildfire occurrences by Distance**

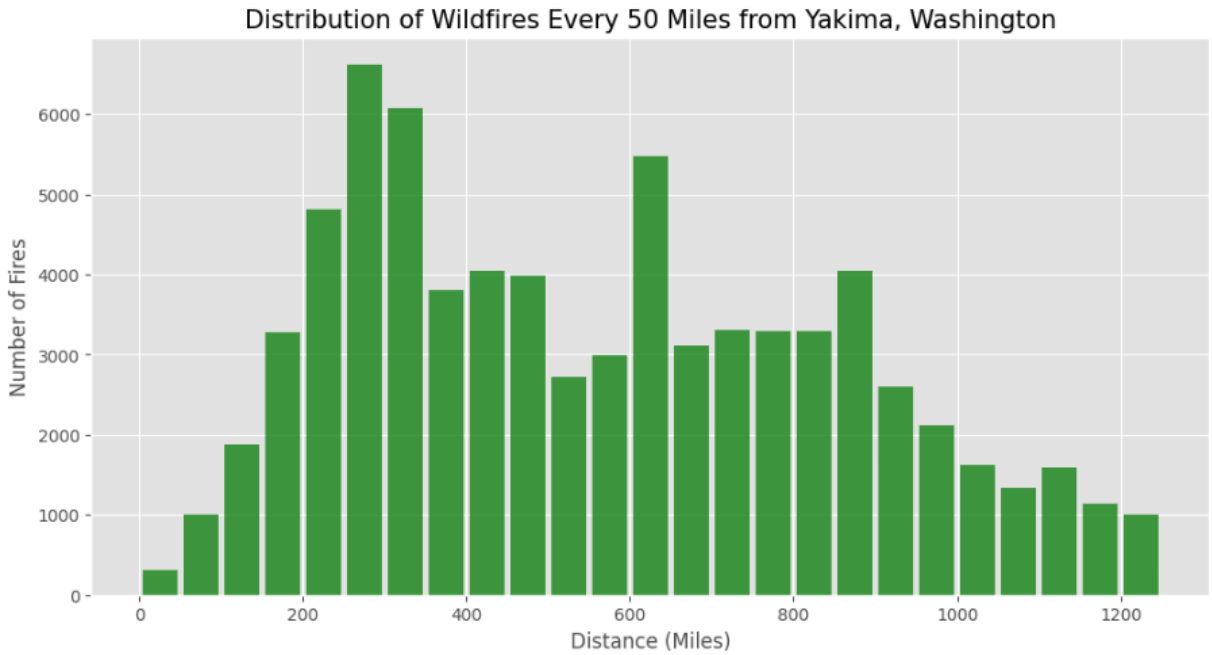


Figure1: A graph of the number of wildfires every 50 miles from Yakima, Washington.

The Figure1's histogram displays the frequency of wildfire occurrences in 50-mile increments from Yakima, Washington, based on data that likely includes fire dates, distances from the city, smoke estimates, areas burned, fire types, and other parameters. The x-axis details the distance in miles, while the y-axis indicates the number of fires. The chart reveals a widespread distribution of fires, with a higher frequency of fires occurring further from Yakima. The peak of fire occurrences falls between distances of 400 to 450 miles, indicating that the immediate vicinity of Yakima has fewer fires. This visual suggests that Yakima may not be a major center for wildfires, as the number of fires decreases closer to the city.

b. Time Series of Total Acres Burned Per Year

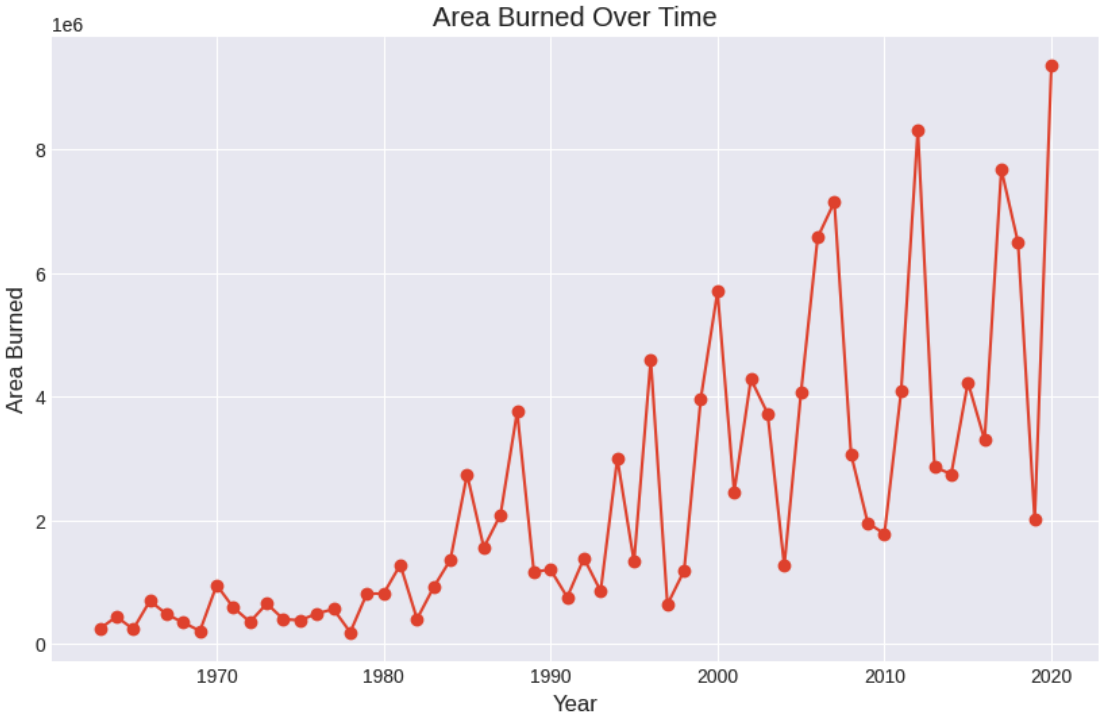


Figure 2: Area Burned over time for Yakima, Washington

The Figure 2's graph shows the annual area burned by wildfires from the 1970s to the 2020s. It highlights a general increase in the extent of wildfires over time, with significant variability from year to year. The peaks indicate years with extensive wildfires, suggesting more severe fire seasons, while the valleys suggest less activity. The upward trend in recent decades may point to factors such as climate change influencing wildfire behavior.

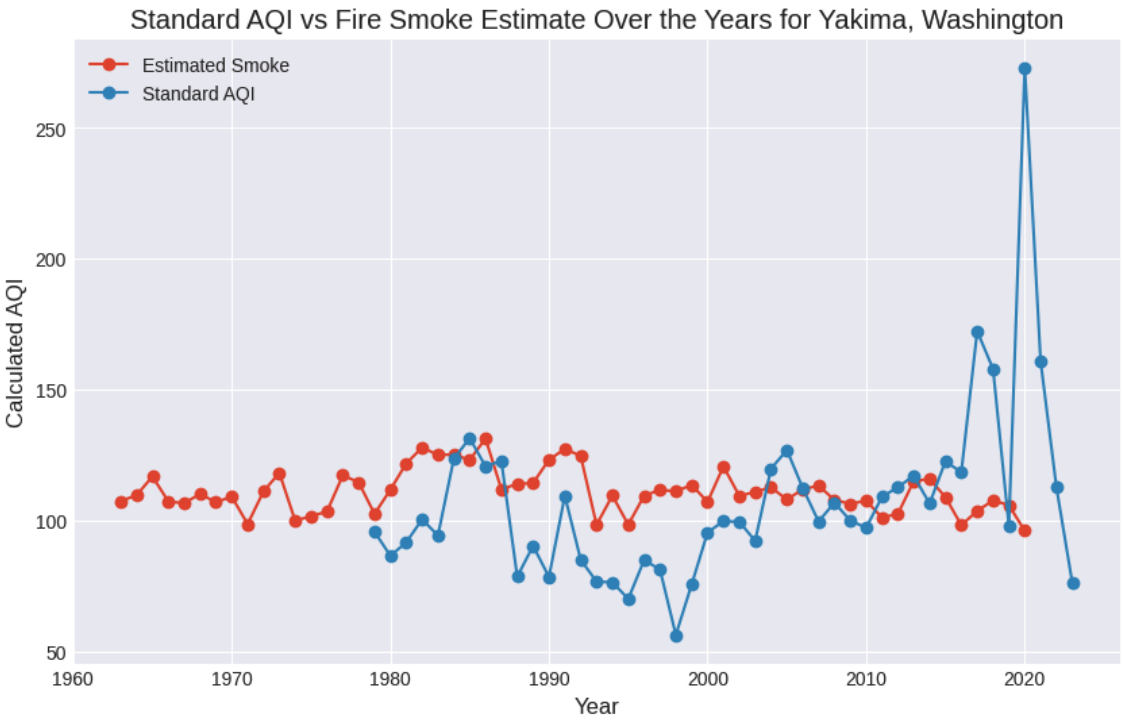


Figure 3: A time series graph containing fire smoke estimate for Yakima city and the AQI estimate for Yakima county.

Figure 3 compares the Standard Air Quality Index (AQI) to an estimate of smoke from fires over the years in Yakima, Washington. The red line represents the estimated smoke, and the blue line shows the standard AQI values.

From the 1960s to the late 2010s, both the estimated smoke and the standard AQI fluctuate, but they generally track closely together, suggesting that the presence of smoke from fires is a significant factor in the overall air quality of the area. However, from around 2015 onwards, there is a notable divergence where the peaks in estimated smoke dramatically exceed the standard AQI values, indicating episodes of intense fire activity with severe impacts on air quality. The largest spike in estimated smoke occurs just before 2020, reflecting an exceptionally severe fire season.

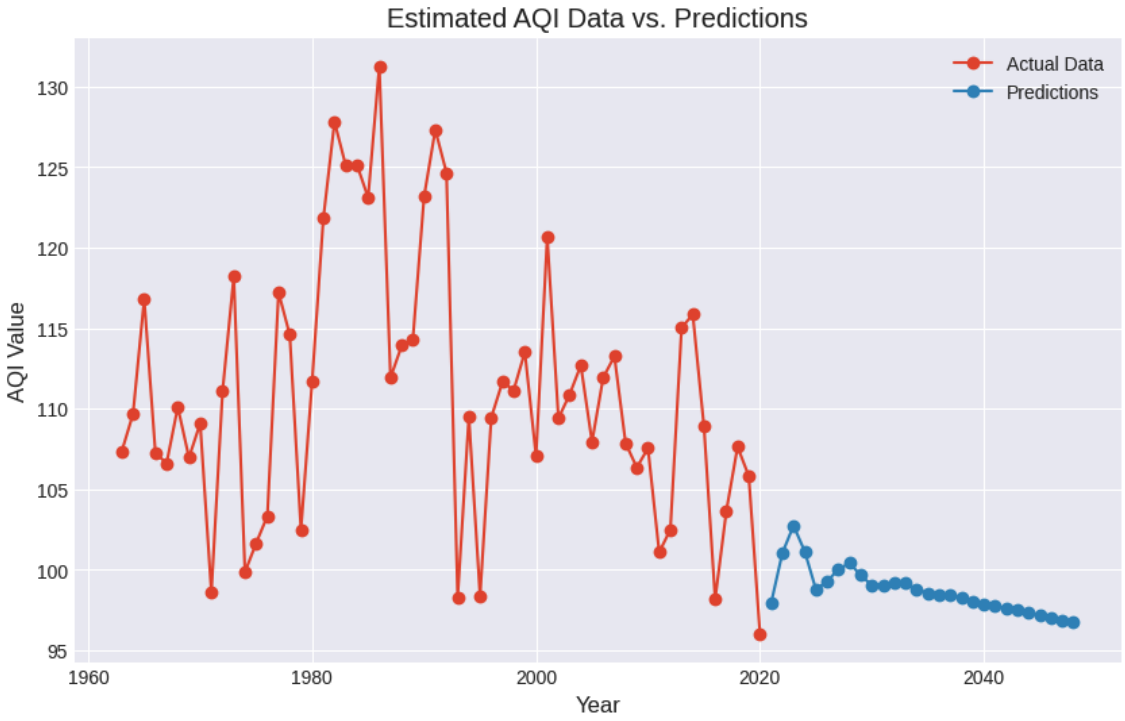


Figure 4: Using ARIMA for Smoke Estimation

The graph presents a comparative view of estimated Air Quality Index (AQI) values against predicted AQI over several decades, beginning from 1960 through to a forecast ending in 2040. The red line denotes the actual AQI data, characterized by fluctuating values that exhibit some form of cyclical pattern, potentially corresponding to the seasonal nature of wildfires and their varying intensities over the years. In contrast, the blue line represents predicted future AQI values, which show a notable downward trend as we move into the future. This decline suggests an anticipated improvement in air quality or a reduction in wildfire-induced pollution, possibly due to projected changes in environmental policies, preventive measures, or other mitigating factors. The graph serves as an analytical tool to assess past air quality challenges and plan for a healthier environmental future based on the predictions made.

2. Hypothesis 2: Correlation between Smoke and Asthma Rate amongst young population in Yakima
    a. Correlation between Asthma Rate and Smoke Estimate

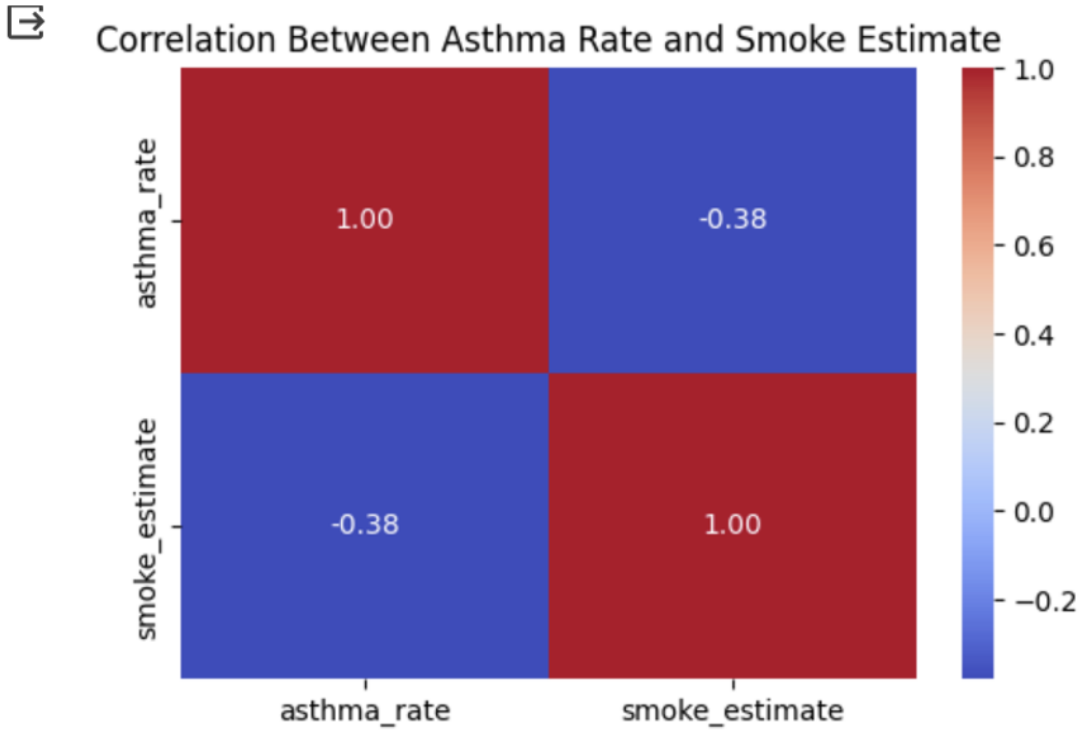


Figure 5: Correlation analysis between Asthma Rate and Smoke Estimate

The correlation analysis between asthma rates and smoke estimates as shown in Figure 5 is a statistical approach to understand the relationship between air pollution caused by wildfires and the prevalence of asthma. The heatmap displayed, with a correlation coefficient of -0.38, suggests a moderate inverse relationship; as smoke estimates increase, asthma rates tend to decrease, or vice versa. This could be due to a variety of factors, including population displacement during heavy smoke periods, underreporting of asthma cases, or other environmental variables not accounted for in the analysis. The findings call for a deeper investigation into the environmental health dynamics affecting Yakima's population, particularly its young adults. It underscores the importance of rigorous air quality monitoring and health surveillance systems to develop effective public health responses to mitigate the adverse effects of wildfire smoke.

b. Asthma Rate Trends Over Time with Smoke Estimate

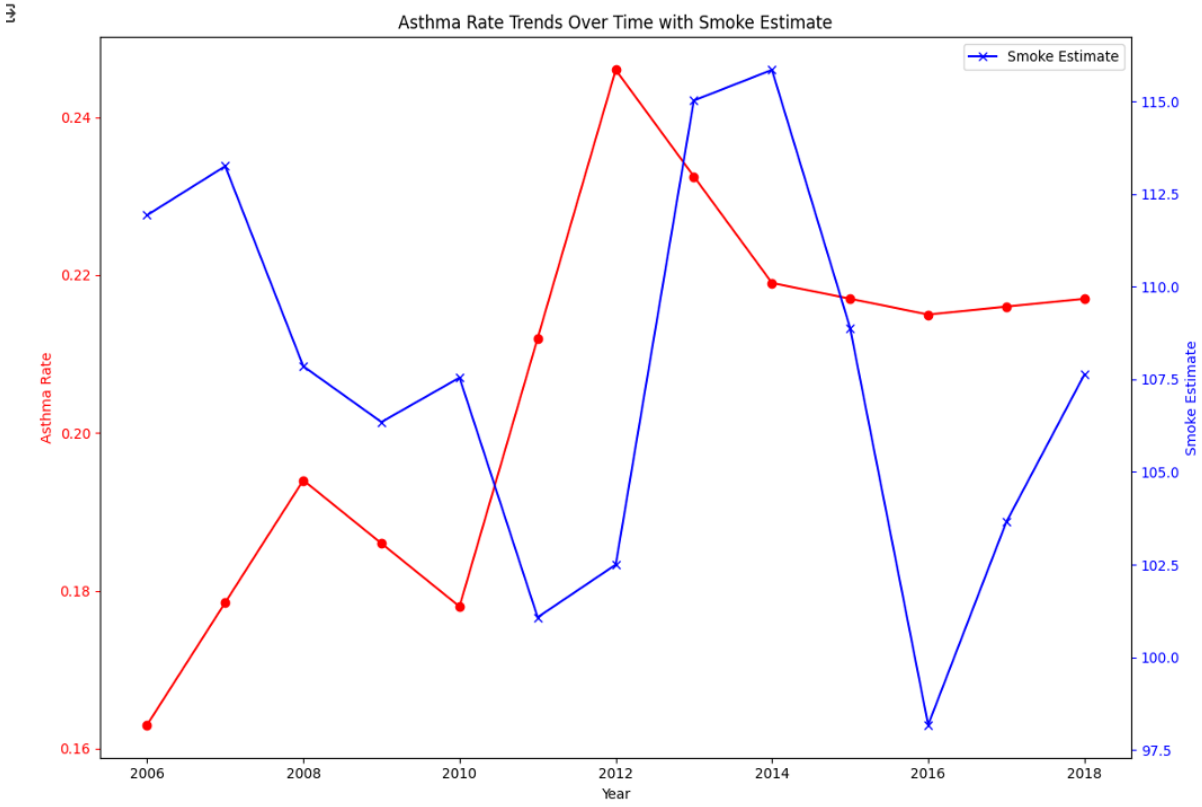
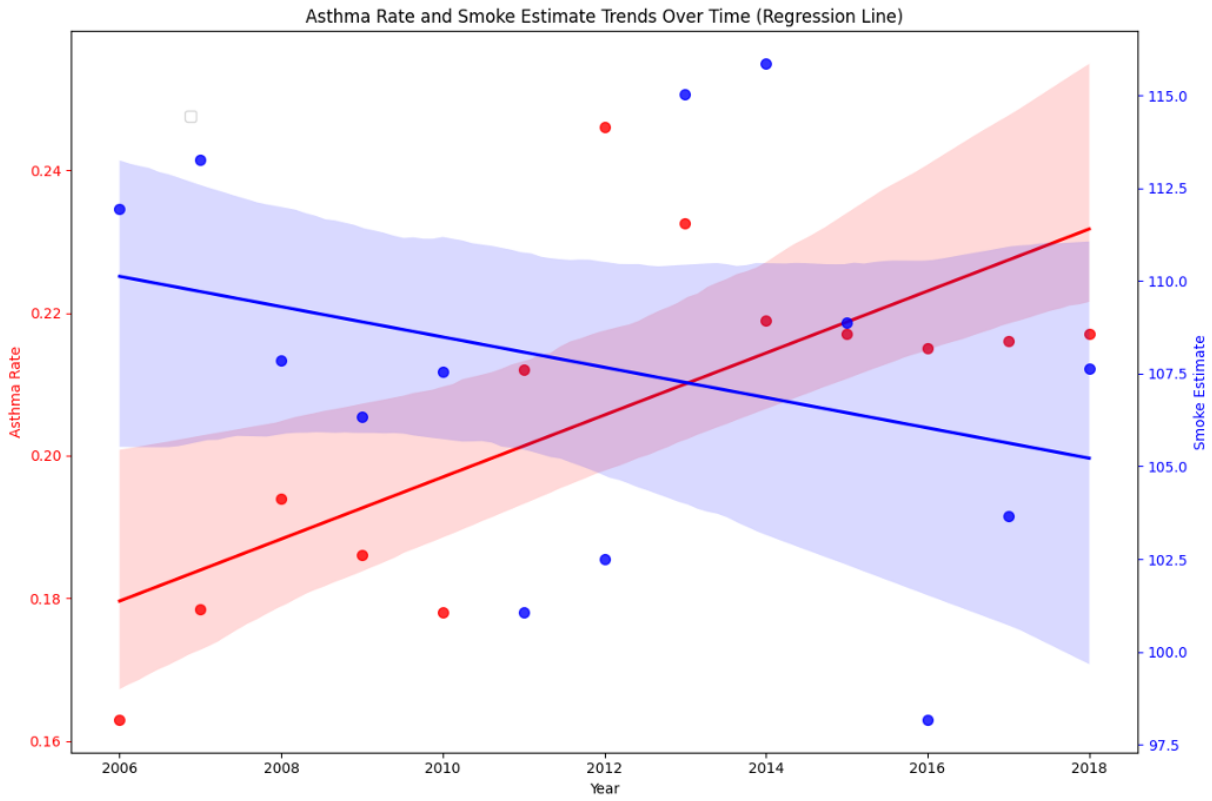


Figure 6: Asthma Rate Trends Over Time with Smoke Estimate

Figure 6 contrasts the actual Air Quality Index (AQI) values over time with predicted values extending into the future. The red line indicates the historical AQI data, showing fluctuations over the years. There's a notable trend of variability, with some peaks suggesting periods of poorer air quality. The blue line represents the forecasted AQI, which appears to decline steadily over time. This predictive element suggests an expected improvement in air quality or a reduction in factors contributing to AQI, such as wildfire smoke. The graph serves as a tool for visualizing past AQI trends and anticipating future conditions, which can be crucial for environmental planning and public health strategies.

c. Trends over time with the Regression Line



The graph showcases the correlation between asthma rates and smoke estimates, reflected by a negative correlation coefficient of -0.38. This indicates an inverse relationship, where higher smoke estimates from wildfires might be associated with lower asthma rates, or vice versa. However, the correlation is not strong enough to imply a direct causation. The data points, represented by dots, are overlaid with a regression line (and its confidence interval) for each variable, displaying the trend over time. The regression lines cross, suggesting changing relationships between the two variables throughout the years, further implying that other factors could be influencing the asthma rates beyond just the smoke estimates. The shaded areas around the regression lines indicate the confidence intervals, which show the range of values within which we can expect to find the true regression line with a certain level of confidence, usually 95%. The graph aims to explore complex environmental health dynamics and serves as a starting point for deeper investigation into the multifaceted interactions between air quality and public health.

DISCUSSIONS

Public Health and Safety: My findings highlight a direct link between wildfire smoke and air quality, which has significant implications for public health, particularly concerning asthma rates among the young population in Yakima. Understanding these relationships is crucial for developing strategies to mitigate health risks.

Environmental Management: The increasing trend of wildfire occurrences and their impact on air quality underscore the need for effective environmental management and proactive measures to address these challenges.

Recommendations for City Council, City Manager/Mayor, and Residents

1. Developing a Comprehensive Wildfire Response Plan: Based on the findings, it's imperative for the city council and mayor to formulate a concrete plan within the next 12 months, before the onset of another wildfire season. This plan should include strategies for air quality monitoring, healthcare preparedness, and community education on wildfire smoke.
2. Enhanced Air Quality Monitoring: Implementing advanced monitoring systems for air quality can provide real-time data, helping to predict and mitigate the effects of smoke on public health.
3. Healthcare Services Expansion: Expanding healthcare services, particularly for respiratory issues, is essential. This includes providing resources and support for asthma patients, especially during high-risk periods.
4. Community Education and Preparedness: Conducting educational campaigns to raise public awareness about the health risks associated with wildfire smoke and ways to reduce exposure is crucial.
5. Environmental Policies and Green Initiatives: Encouraging policies that focus on reducing wildfire occurrences through environmental conservation and green initiatives.

Human-Centered Data Science Reflection

1. Contextual Understanding: The project's approach was deeply rooted in understanding the unique environmental and climatic context of Yakima. This enabled a focused analysis that is relevant to the specific challenges faced by the city.

2. Stakeholder Engagement: The analysis and recommendations were shaped by considering the needs and roles of various stakeholders, including local authorities, healthcare providers, and residents. This ensured that the solutions proposed are practical and targeted.
3. Public Health Integration: By linking environmental factors with public health concerns, the project emphasizes the need for an integrated approach that addresses both ecological balance and the well-being of the community.
4. Adaptive Strategies: The recommendations are not static; they are designed to be adaptable to changing conditions and new data. This aligns with the dynamic nature of environmental challenges and the need for continuous reassessment and adjustment of strategies.

LIMITATIONS

The investigation into the effects of wildfire smoke on asthma rates in Yakima, using ARIMA modeling, faces inherent limitations that must be acknowledged for a comprehensive understanding of the study's scope. The expansive radius for data collection may dilute the specific impact on Yakima, while the smoke estimate, lacking in environmental variables such as wind patterns and topography, might not capture the true extent of exposure. The AQI data's exclusion of gaseous pollutants and reliance on remote weather stations could skew the representation of local air quality. Seasonal variations in wildfires, a crucial aspect of accurate modeling, were not accounted for due to the absence of granular monthly data.

The handling of the vast and variable GeoJSON dataset presents its own set of analytical challenges. The ARIMA model, while adept at capturing trends, may inadvertently apply historical trends to future predictions, which may not hold true given the formulaic basis of the smoke estimate. The focus on asthma, excluding other respiratory conditions potentially influenced by smoke, narrows the study's health implications. Furthermore, the application of linear regression without verifying its assumptions may affect the robustness of the findings. Despite these limitations, the study's commitment to a human-centered design ensures that the outcomes remain focused on the welfare of Yakima's community, striving to inform and support local decision-making.

CONCLUSIONS

This comprehensive study on "Understanding the Impact of Wildfires on Yakima, Washington" has illuminated the multifaceted and dynamic nature of wildfires, particularly their significant impacts on air quality and public health. Through meticulous analysis and methodical research, the project has shed light on the intricate connections between environmental factors and health outcomes, specifically focusing on the prevalence of asthma among the youth in the wake of increasing wildfire smoke.

Key Contributions and Insights:

- The study's exploration of wildfire occurrences and their distribution relative to Yakima provided critical insights into regional wildfire patterns, informing local preparedness and response strategies.
- The correlation analysis between smoke estimates and asthma rates, despite presenting a moderate inverse relationship, underscored the necessity of nuanced interpretations and deeper investigation into environmental health dynamics.
- Predictive models, such as ARIMA, employed in the study, offered a forward-looking perspective on air quality trends, presenting a valuable tool for urban planning and public health forecasting.

Human-Centered Data Science Perspectives

1. Understanding Human Impact: The study underscores the essence of human-centered data science by focusing on the human impact of environmental phenomena. It goes beyond mere statistical analysis to interpret how wildfire smoke affects people's health, particularly the young population, thereby grounding data science in real-world human experiences.

2. Informative and Actionable Insights: By translating complex environmental data into understandable and actionable insights, the study serves as a guide for policymakers, healthcare providers, and the community in formulating effective response strategies. It exemplifies the role of data science in informing decisions that directly impact human well-being.
3. Stakeholder Engagement and Ethical Considerations: The research was conducted with a strong emphasis on stakeholder engagement, considering the needs, challenges, and roles of various community members. Ethical considerations were paramount, focusing on the welfare and health of the Yakima community, thereby reflecting the core principles of human-centered data science.
4. Adaptability and Limitations: Acknowledging and addressing the limitations of the study, such as potential data biases and the specificity of methodologies, demonstrates the adaptive nature of human-centered data science. It highlights the importance of continually refining methods and models in response to new data and changing circumstances.
5. Empowering Communities: By providing insights into the correlation between environmental factors and health outcomes, the study empowers communities to advocate for and engage in informed decision-making. It reinforces the idea that data science should not just be about data, but about how data can be used to improve human lives.

REFERENCES

1. Wildfire Polygons Metadata Explanation: https://www.sciencebase.gov/catalog/file/get/61aa537dd34eb622f699df81?f=__disk__d0%2F63%2F53%2Fd063532049be8e1bc83d1d3047b4df1a5cb56f15&transform=1&allowOpen=true
2. Wildfire GeoJSON module (Prof. David McDonald) https://drive.google.com/file/d/1TwCkvdaw0MxJzW7NSDg6XxYQ0dvaS44I/view
3. AQI Metadata Explanation: https://www.airnow.gov/sites/default/files/2020-05/aqi-technical-assistance-document-sept2018.pdf

4. Asthma Data for 10th graders https://datacenter.aecf.org/data/tables/5072-percentage-of-tenth-graders-with-asthma?loc=49&loct=5#detailed/5/6947-6985/false/37,870,869,868,133,35,17/any/11494
5. ARIMA time series forecasting https://www.kdnuggets.com/2023/08/times-series-analysis-arima-models-python.html#:~:text=The%20ARIMA%20model%20is%20a,making%20skillful%20time%20series%20forecasts.

DATA SOURCES

1. Historical Wildfire Occurrences Dataset: https://www.sciencebase.gov/catalog/item/61aa537dd34eb622f699df81
2. US EPA API - Air Quality IndexDataset: https://aqs.epa.gov/aqsweb/documents/data_api.html
3. Asthma Data for 10th graders https://datacenter.aecf.org/data/tables/5072-percentage-of-tenth-graders-with-asthma?loc=49&loct=5#detailed/5/6947-6985/false/37,870,869,868,133,35,17/any/11494