DATA 512 Human Centered Data Science

12/4/2024

Kelly Wang

Economic Impacts of Wildfire Smoke in Chico, CA

Introduction

Wildfires are a growing concern for many communities across California, including Chico in Butte County, which experienced the Park Fire this year, California's fourth-largest wildfire on record, as well as the Camp Fire in 2018, California's most destructive wildfire to date. Chico is located in the northern central part of California, and it is the most populous city in Butte County.

The frequency and intensity of wildfires in California have been increasing due to a combination of factors such as climate change and drought. These wildfires pose substantial threats to the county's public health, local economics, service industry, education, community differences, and overall well-being. By quantifying the economic impacts of wildfire smoke, this analysis aims to provide critical insights that can inform policymakers, city managers, city councils, or other stakeholders in developing effective mitigation and recovery strategies. These insights can help protect jobs, support industries that are most vulnerable to wildfires, and better prepare to enhance the overall well-being of the residents.

Background/Related Work

Background Research

Research on the economic impact of wildfires reveals significant shifts in Butte County's economy. Specifically, analysis by the Northern Rural Training and Employment Consortium (NoRTEC) on Butte County's labor market and industry sector found that the Camp Fire in November 2018 led to drastic loss of life, property and income. However, spending on suppression and recovery efforts provided short-term economic growth. In the first six months after the Camp Fire, Butte County experienced "a net positive economic effect, with employment increasing by 0.6%, earnings by 1.8%, and economic output by 0.7%" (Northern Rural Training and Employment Consortium [NoRTEC], 2019).

Research Questions

In this analysis, we build on the prior background research and explore how wildfire smoke affects employment and economic outcomes, focusing on primary industries in Butte County over the past years. I hope to uncover patterns in employment that correlate with wildfire events by examining historical employment and GDP data in Butte County.

Specifically, the following research questions will be investigated:

1. What economic impact does wildfire smoke have on Butte County's primary industries?
2. Does wildfire smoke significantly affect the city's employment rates and GDP?

My hypothesis is that there could be a relationship between the increasing smoke impacts over the past years and Butte County's economic indicators. However, I am curious whether the damage caused by smoke outweighs the economic benefits brought by recovery efforts. I hypothesize that the major wildfire events and the resulting smoke have notable effects on employment rates and GDP growth. Understanding these relationships will help inform policy and support strategies for policymakers in Butte County.

Data

Wildfire Data: The wildfire data is based on the Combined wildland fire datasets for the United States and certain territories, 1800s-Present (combined wildland fire polygons), collected and aggregated by the US Geological Survey. The dataset provides fire polygons in ArcGIS and GeoJSON formats. Specifically, the combined data stored in GeoJSON format as "USGS_Wildland_Fire_Combined_Dataset.json" is used. As a result of the data acquisition and preparation process, I accessed 117,190 wildfire events for the city of Chico after filtering to include only the last 60 years of wildland fire data (1964-2024).

AQI Data: To evaluate the quality of the smoke impact estimates, Air Quality Index (AQI) data is used in this analysis and retrieved from the US Environmental Protection Agency (EPA) using the US EPA Air Quality System (AQS) API. This API provides historical air quality data but does not offer real-time information. The documentation outlines the different call parameters and provides examples of the various calls that can be made to the API.

The EPA reports that broad-based monitoring with standardized quality assurance procedures began in the 1980s, but some counties still lack air quality monitoring stations. I retrieved AQI data specific to Chico, CA, by querying the API with Chico's FIPS code, which helped identify the monitoring station located at "San Andreas-Gold Strike Road." However, AQI data for Chico, California, is only available from 1994 to 2020. Additional information can be found in the EPA's Frequent Questions about AirData page, terms of use of the API can be found in this section, and explanation of how AQI is calculated from pollutant concentration can be found in the AQI Technical Assistance Document.

In addition to the wildfire dataset collected by the US Geological Survey and the Air Quality Index (AQI) data, I incorporated two data sources related to the economic indicators of Butte County:

1. Employment: The California monthly employment data by industry (Current Employment Statistics (CES))
2. GDP: California Gross Domestic Product by County (California Regional Economic Analysis Project)

Employment Data: The employment data is sourced from the Current Employment Statistics (CES) program, a Federal-State cooperative effort. The program conducts monthly surveys to provide estimates of employment, hours, and earnings based on payroll records of business establishments. The data reflects the number of nonfarm, payroll jobs and it excludes proprietors, self-employed, unpaid family or volunteer workers, farm workers, and household workers. The dataset is publicly available and published under the Creative Commons Attribution.

The dataset consists of employment statistics by industry for various regions in California, from 1990 to 2024. Each employment record is represented by key columns, including a unique identifier, geographical classification, area name, year, month, date, industry title, and the number of individuals employed in each industry for that specific month. This employment data is a valuable resource for analyzing employment trends over time and understanding how different industries are affected by factors like wildfire. By leveraging the CES data, I hope to analyze historical employment trends in Butte County, identify the industries most impacted by wildfires in terms of employment, and find correlations between wildfire events and employment changes.

GDP Data: The GDP data is sourced from the California Gross Domestic Product by County (Butte County), provided by the California Regional Economic Analysis Project (CA REAP). This data offers annual estimates of economic output at the county level, capturing the value of goods and services produced. The dataset includes information on various sectors, such as agriculture, manufacturing, and services, offering a comprehensive view of economic activity across California's counties.

The dataset spans from 2001 to 2022 and includes key columns such as year, industry title, and the gross domestic product (GDP) value. With the CA REAP GDP data, I aim to examine the historical economic trends in Butte County, identify the industries most impacted by wildfires in terms of GDP, and explore the correlations between wildfire events and changes in economic output.

Methodology

This section outlines detailed methodology used for the entire analysis, including how the initial smoke impact estimate was quantified and how the economic impact was analyzed. Additionally, ethical considerations and the rationale for using these methods are explained.

Estimating Wildfire Smoke

Understanding the amount of smoke exposure from wildfires is crucial for identifying the relationships between smoke impacts and economic outcomes, as well as for gaining a better understanding of their implications for residents. To do this, I filtered the wildfire data to include only fires that occurred within the past 60 years (1964-2024). For each wildfire event, the fire year, name, size (measured in acres), type, and distance from Chico city were considered. For this analysis, I also filtered the wildfire data to include fires that occurred within 650 miles of Chico. After all the filtering steps, 55,153 wildfires remain from the initial set of 117,190 wildfire events.

To estimate the smoke impact, we typically rely on data that includes PM2.5, which is fine particulate matter measured in micrograms per cubic meter, or AQI (air quality index) data produced by EPA (Washington State Department of Labor & Industries, n.d.). However, for this analysis, I only have access to limited wildfire-related data, such as the fire size, fire type, and distance to the city. To estimate smoke exposure based on the characteristics of the wildfire without direct AQI or PM2.5 data, I used information on fire type, area burned, and distance to Chico. Since the fire type variable is categorical, I assigned priority scores to wildfires based on their potential impact on smoke emissions, ranked by fire type. These priority scores, or fire type ranking weights, help quantify each fire type's contribution to the smoke estimate. A higher weight corresponds to greater anticipated smoke emissions from that fire type. For example, the "Wildfire" category received the highest weight, indicating the highest potential for smoke emissions, while "Unknown - Likely Prescribed Fire" received the lowest weight, reflecting a lower expected smoke impact. Based on these factors, the smoke estimate was calculated for each wildfire event. The formula used is as follows:

$$\text{smoke\_estimate} = \frac{\text{area\_burned} \times \text{fire\_type\_weight}}{\text{distance}}$$

Figure 1: Formula used to calculate smoke estimate

Although this calculation logic does not directly measure PM2.5 concentrations, it provides a reasonable estimate of smoke exposure based on available wildfire data. This approach ensures that larger, closer, and more severe fires are assigned higher smoke estimate values.

The smoke estimate results were validated by comparing with the AQI data during the wildfire season. I combined daily AQI data from gaseous and particulate pollutants, aggregated it annually, and calculated an average AQI per year to estimate wildfire smoke's impact to compare with my smoke estimates. Because AQI data is available only from 1994 to 2020, covering less than half of the 60-year wildfire data span. Therefore, estimating smoke exposure using the above formula using wildfire characteristics helps fill the gaps in AQI data.

Analyzing Economic Impacts

Now that the smoke estimate was calculated for each year, I moved on to investigate the economic impact wildfire smoke has on Butte County’s primary industries. To do this, the methodology includes integrating GDP, employment, and smoke estimate data, identifying the top 10 primary industries in Butte County, and analyzing correlations between wildfire smoke estimates and economic indicators of these industries to uncover potential relationships. As wildfire impacts on the economy might not be immediate. I then explored lagged effects where employment impacts might occur 1 year after the wildfire events.

Human-centered Considerations

In designing this study, I aimed to balance the analytical methods with a human-centered approach that emphasizes ethics and community impact. The decision to estimate smoke exposure using wildfire characteristics was driven by the need to address critical data gaps while maintaining accuracy through validation with AQI data. By prioritizing publicly available data from trusted sources, I ensured that the analysis used datasets under transparent licensing agreements and excluded personally identifiable information through aggregation or anonymization. Additionally, the work is fully reproducible, with detailed documentation and commit history available in my Git repository.

Findings

Wildfire and Smoke Patterns

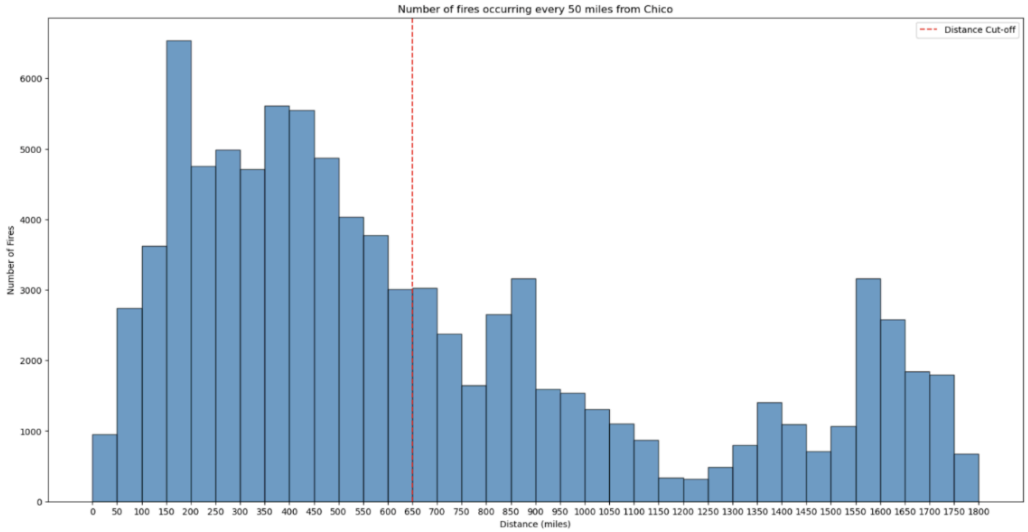

Figure 2: Wildfire Frequency near Chico, CA

Figure 2 displays the number of fires occurring at 50-mile intervals from Chico, for all fires up to 1,800 miles away. The distance cutoff at 650 miles represents the data included in the smoke estimate. From this histogram, we can see that wildfires appear more frequently in areas closer to Chico, with the highest number occurring within the 500 miles around Chico.

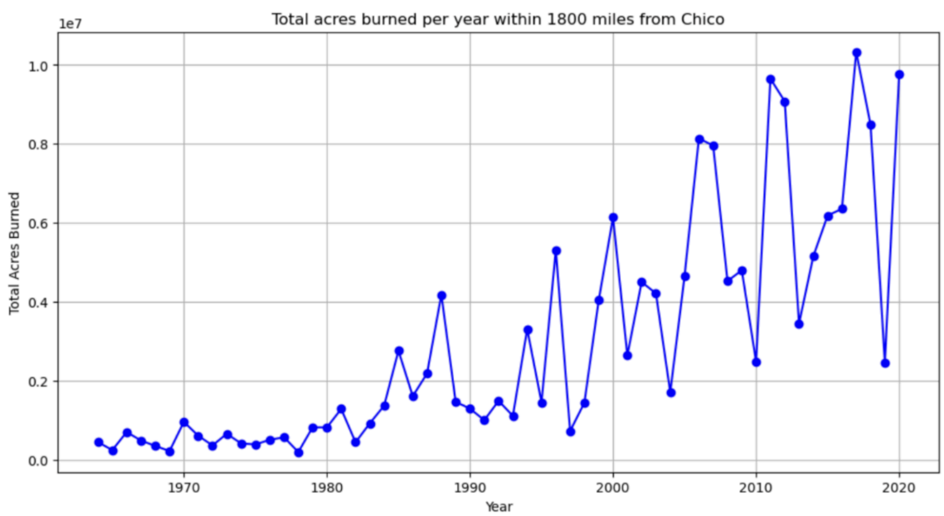

Figure 3: Annual Total Acres Burned by Wildfires near Chico, CA

Figure 3 shows a time series of the total acres burned per year for wildfires occurring within 1800 miles from Chico, CA, over the last 60 years. The x-axis represents the years and the y-axis represents the total acres burned per year for the wildfires occurring within 1800 miles. The time period for the data spans from 1964 to 2024.

From this time series plot, we can see the annual total acres burned near Chico generally increased over time. However, it should also be noted that this increase could be partially attributed to improved observation and recording capabilities enabled by advances in technology.

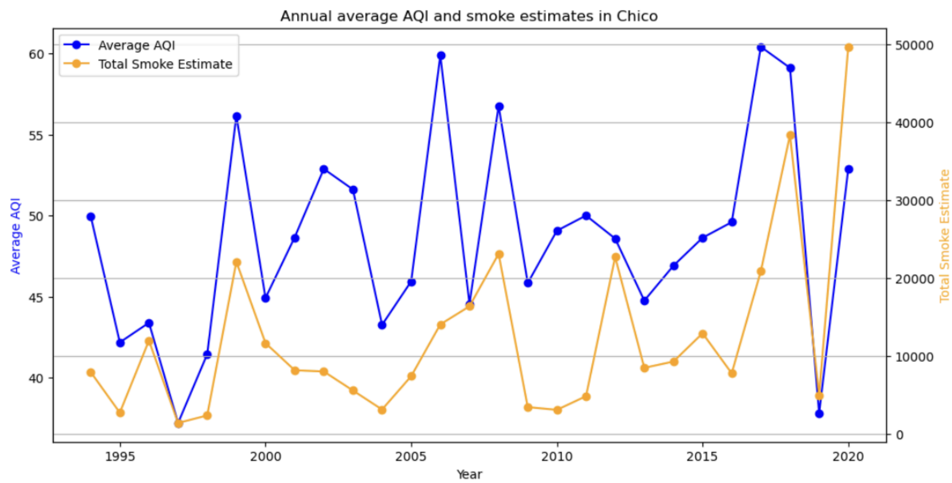

Figure 4: Annual Average AQI and Annual Smoke Estimates Comparison

Figure 4 compares the annual average AQI and the annual total fire smoke estimates for Chico, CA, over the last 60 years (1964 to 2024). The left y-axis represents the annual average AQI, plotted in blue, while the right y-axis represents the total smoke estimates, plotted in yellow. We can observe that there are certain years where both the AQI and smoke estimates spike together (for example, 1999, 2008, and 2018), suggesting a possible positive correlation between AQI and smoke impact. This correlation indicates that the smoke estimate is reasonably accurate, as it aligns well with the observed AQI data.

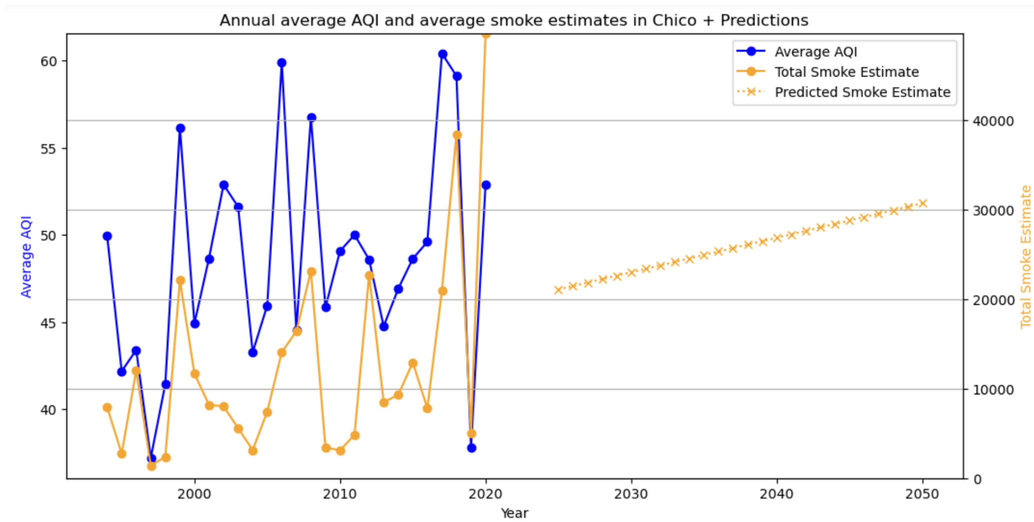

Figure 5: Annual Average AQI and Annual Smoke Estimates Comparison with Predicted Smoke Estimates

After developing a predictive model using linear regression based on the fire data and smoke estimate for Chico, I predicted smoke estimates for every year for the next 25 years (2025-2050), as shown in **Figure 5**. The predicted smoke estimates, represented by the orange dashed line, show a clear upward trend. This suggests an increasing impact of wildfire smoke over time.

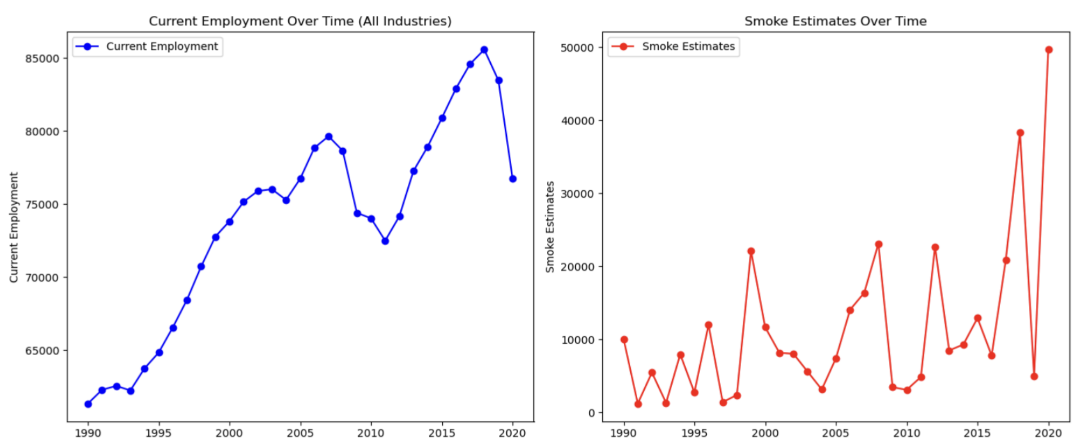

Figure 6: Employment and Smoke Estimate Trend

Figure 6 shows that there is an overall increasing trend in employment across all industries from 1990 until around 2007. After 2007, the employment trend became more unstable, with noticeable declines around 2011 and 2020. In comparison, the smoke estimates show significant variability, with a high degree of fluctuation over the years. There are notable peaks around the years 1999, 2008, 2012, 2018, and 2020, which might correspond to major wildfire events such as the 2018 Camp Fire in Northern California, and the 2020 North Complex Fire.

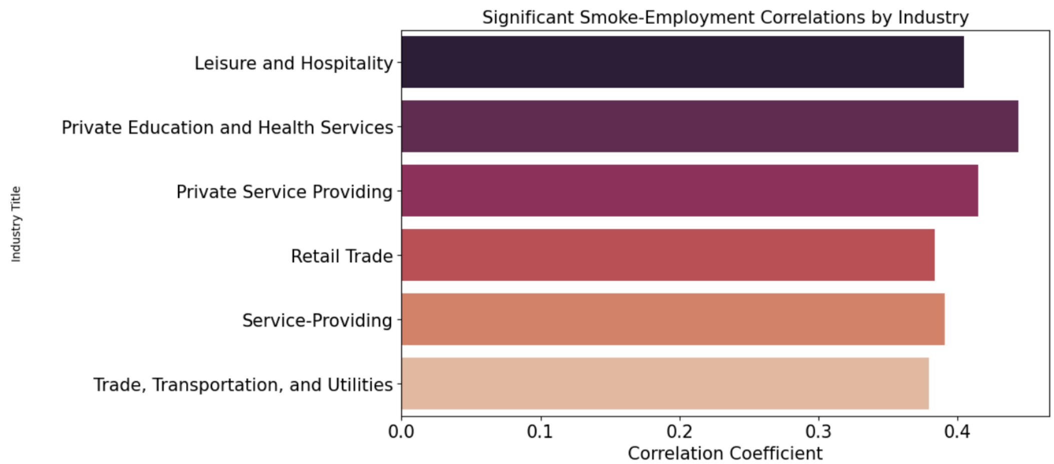

Figure 7: Significant Smoke-Employment Correlations by Industry

From the results shown in **Figure 7**, there is a moderate positive correlation between smoke estimates and employment across various industries. Leisure and Hospitality, Private Education and Health Services, Private Service Providing, Retail Trade, Service-Providing, and Trade, Transportation, and Utilities all show correlation coefficients around 0.4. Interestingly, this indicates that increased smoke estimates generally correspond to higher employment levels in these sectors.

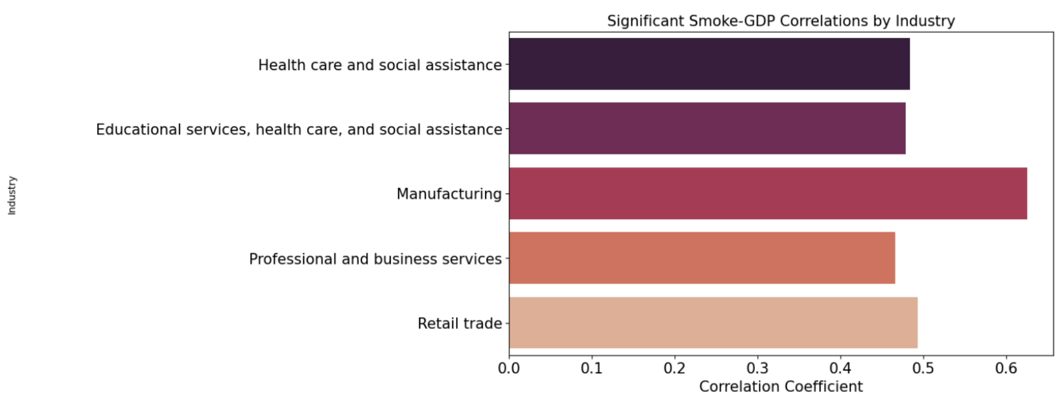

Figure 8: Significant Smoke-GDP Correlations by Industry

Using the GDP data and following a similar methodology, there are noticeable positive correlations between smoke estimates and GDP across several industries as well, as shown in Figure 8. The industries with significant moderate correlations include Health care and social assistance, Educational services, Manufacturing, Professional and business services, and Retail trade. These industries exhibit correlation coefficients around 0.5, similar to the employment correlations. This suggests that higher smoke estimates are associated with increased GDPs in these sectors.

The lag analysis reveals that the impacts of wildfires on employment might not be immediate, with moderate positive correlations potentially occurring one year after the wildfire events for certain industries. This suggests that higher smoke estimates are associated with increased employment in these industries (Leisure and Hospitality, Private Education and Health Services, Government, Private Service Providing, Retail Trade) one year later, indicating potential delayed economic responses to wildfire events.

The lag analysis for GDP reveals that Professional and Business Services and Private Industries show moderate positive correlations between smoke estimates and GDP one year after wildfire events. This indicates that increased smoke estimates are associated with higher GDP in these sectors one year later.

Discussion/Implications

The positive economic growth could indicate increased demand for services related to recovery efforts and population relocation. As a result, certain industries may experience economic benefits following wildfire events, despite the immediate negative impacts of smoke exposure. For example, after major wildfires, displaced individuals and families require housing, healthcare, social assistance, and educational services, all of which likely contribute to short-term job creation and increased economic activity in Butte County. However, these economic benefits may be short-lived and do not necessarily reflect the long-term health and environmental costs of wildfire smoke. The economic growth associated with wildfire recovery efforts, such as increased employment in construction, temporary housing, and local services, may be overshadowed by the longer-term costs of healthcare expenditures and productivity losses due to poor air quality.

The predictive models from this analysis suggest an increasing trend in wildfire smoke exposure over the next 25 years, which adds urgency to addressing these challenges. Given this trend, it is critical that policymakers take immediate action, particularly in the next 1-5 years. Specifically, policymakers in Butte County should prioritize ensuring adequate housing, educational facilities, and health care and social assistance to support populations affected by wildfire events.

The human-centered principles that informed my decision-making in this project revolve primarily around transparency and accountability. I emphasized transparency in both the methodology and the data sources, ensuring clear communication about how the data was collected and transformed. Accountability in this project was maintained by ensuring that the data, methods, and results were fully traceable, reproducible, and aligned with ethical standards, allowing stakeholders to trust the findings.

Limitations

Every study has limitations, and this analysis is no exception. First, the economic impact analysis is conducted at the county level, using smoke data with fire distance specific to Chico. As a result, the findings still need to be validated at the local level for the city of Chico.

Meanwhile, the Park Fire, which started on July 24 in east Chico, spread rapidly, reaching across 429,603 acres by August 26. At more than 600 square miles, this is 13 times the size of San Francisco, or a larger area than New York City or Los Angeles (Center for Disaster Philanthropy, 2024). The wildfire smoke data I had access to does not include 2024, meaning it does not account for the Park Fire, California's fourth-largest wildfire on record.

Another point is that while the analysis identifies correlations, it does not imply causation. Other underlying factors or confounding variables may influence both wildfire smoke impacts and economic indicators. Additionally, I used GDP and employment as proxies for economic impact. These indicators might not capture all dimensions of economic health, such as income levels, business closures, or lost sales.

The analysis might not fully account for the nuances within an industry category. For example, the retail trade sector might experience both positive and negative impacts from wildfire events, depending on the types of products sold and the adaptability of businesses. Finally, the method used for estimating smoke impacts based on fire type, area burned, and distance may oversimplify the complex dynamics of smoke dispersion and exposure. It assumes that a fire’s smoke generation is only influenced by its size, intensity, and distance to Chico.

Conclusion

This study focused on the economic impact of wildfires on Chico, Butte County, answering two key research questions:

1. What economic impact does wildfire smoke have on Butte County's primary industries?
2. Does wildfire smoke significantly affect the city's employment rates and GDP?

The findings of this study support the hypothesis that increasing wildfire smoke exposure would correlate employment rates and GDP growth. In particular, industries such as Leisure and Hospitality, Private Education and Health Services, Private Service Providing, Retail Trade, Service-Providing, and Trade, Transportation, and Utilities show moderate positive correlations between smoke and employment. The industries such as Health care and social assistance, Educational services, Manufacturing, Professional and business services, and Retail trade also show moderate positive correlations between smoke and

GDP. These results suggest that higher smoke estimates are associated with increased economic activities in these sectors, which could mean wildfire events generate demand for services related to recovery, relocation, and support for displaced populations.

This study exemplifies the principles of human-centered data science by focusing not only on quantitative analysis but also on how the implications and results impact people's lives in Chico, CA. Additionally, this analysis emphasizes transparency by relying on publicly available data sources and being fully reproducible.

References

1. Butte County Labor Market Profile and Industry Sector Analysis by Northern Rural Training and Employment Consortium (NoRTEC) https://www.ncen.org/images/documents/lmi/regional-profiles/2019/Butte%20County%20Report.pdf
2. Washington State Department of Labor & Industries. (n.d.). *Wildfire smoke*. https://www.lni.wa.gov/safety-health/safety-topics/topics/wildfire-smoke#overview
3. *2024 North American wildfires*. Center for Disaster Philanthropy. (2024, October 4). https://disasterphilanthropy.org/disasters/2024-north-american-wildfires/?gad_source=1&gclid=Cj0KCQiAo5u6BhDJARIsAAVoDWtC7wuKR99eR4zitgyG8hQTiYoDF7Lo5ine-Wu2hCZNR9dx01aAqXQaAtWsEALw_wcB

Data Sources

Combined wildland fire datasets for the United States and certain territories, 1800s-Present (combined wildland fire polygons)

US EPA Air Quality System (AQS) API

Current Employment Statistics (CES)

California Gross Domestic Product by County