Analysis and Reflection

Raagul Nagendran

Visualizations

Visualization 1

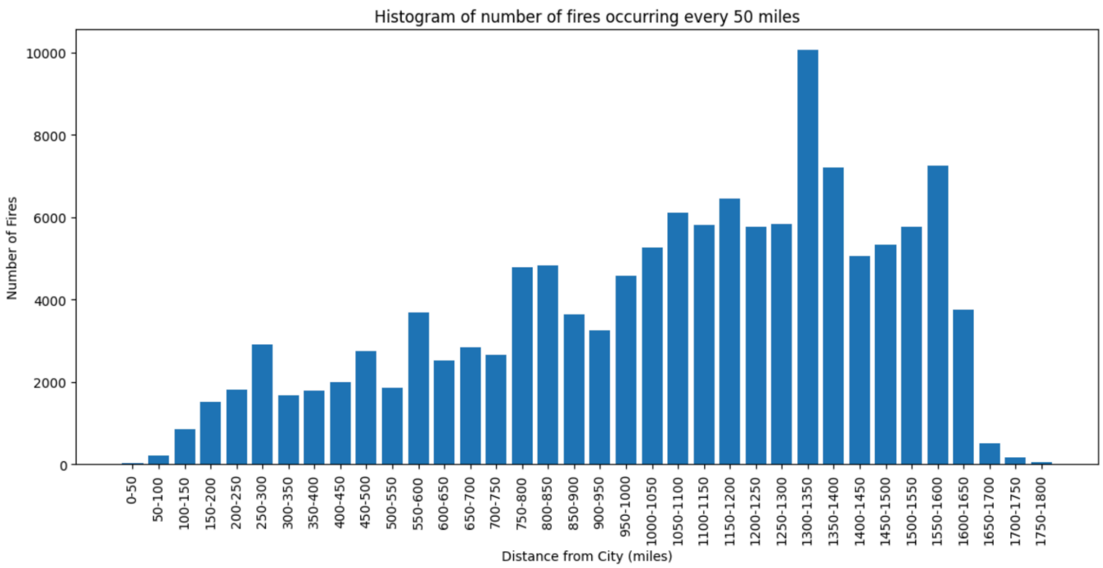

Figure 1: Histogram showing the number of wildfires occurring every 50 miles from Arlington, Texas up to 1800 miles.

The histogram shows the number of fires occurring every 50 mile distance from Arlington, Texas. The x-axis represents the distance from the city in miles, split into bins of 50 miles, up to a maximum of 1800 miles from Arlington. The plot displays a clear pattern, with the number of fires peaking at certain distances from the city and then declining as the distance increases.

The most notable feature of the histogram is the very high number of fires occurring at a distance of around 1300 miles from the city. This peak suggests that there may be a specific region or area located approximately 1300 miles away from the city that is particularly prone to fires. It would be interesting to investigate the underlying factors that contribute to this high concentration of fires at this specific distance, but this is beyond the scope of our current project as our current project is restricted to 650 miles.

Additionally, the plot reveals several other local maxima in the number of fires at various distances from the city, such as around 300 miles, 550 miles, and 700-800 miles. These peaks indicate that there are other regions or factors that may be associated with increased fire occurrences at those distances. Understanding the geographical, environmental, or human-related reasons that lead to these fires could provide valuable insights for fire prevention in the areas surrounding the city.

Visualization 2

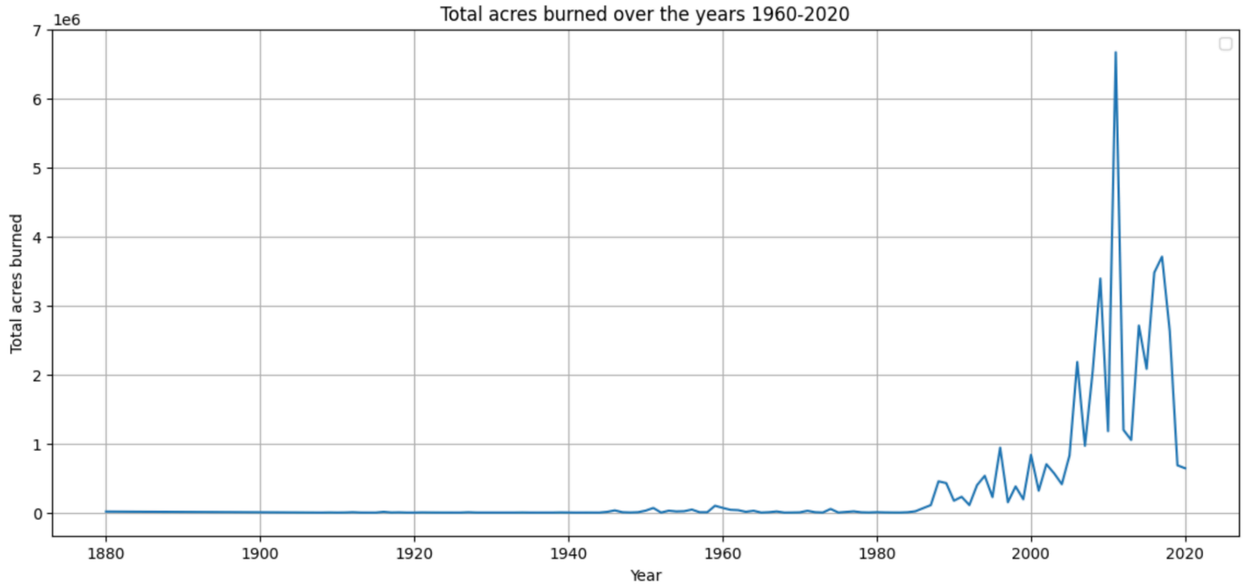

Figure 2: Time series graph of the total acres burnt per year for the fires occurring within 650 miles of Arlington, Texas.

The plot describes the total acres burned over the years from 1960 to 2020. The x-axis shows the years from 1960 to 2020. The y-axis shows the total acres burned in millions of acres (1e6 scale). The image clearly shows a significant increase in the total acres burned during this period, with several notable peaks and valleys starting from 2000.

We notice the steep increase in acres burned starting around the year 2000. This sharp rise might be due to the reason that there have been a lot of changes in the various factors contributing to the frequency and intensity of wildfires in the region over the past two decades. Identifying and understanding the underlying causes of this accelerated trend, such as rising temperatures, changes in land use, drier conditions, or other environmental changes, could provide valuable insights for developing effective strategies to mitigate the impact of these devastating events.

Additionally, the graph reveals several other periods of high fire activity, such as the peaks observed in the 1960s and 1980s. These fluctuations indicate that the region has experienced recurrent challenges with wildfires, albeit with varying magnitudes. Analyzing the factors that may have influenced these historical patterns, as well as any potential shifts in the drivers of fire activity, could help inform long-term planning and preparedness efforts for the region.

Visualization 3

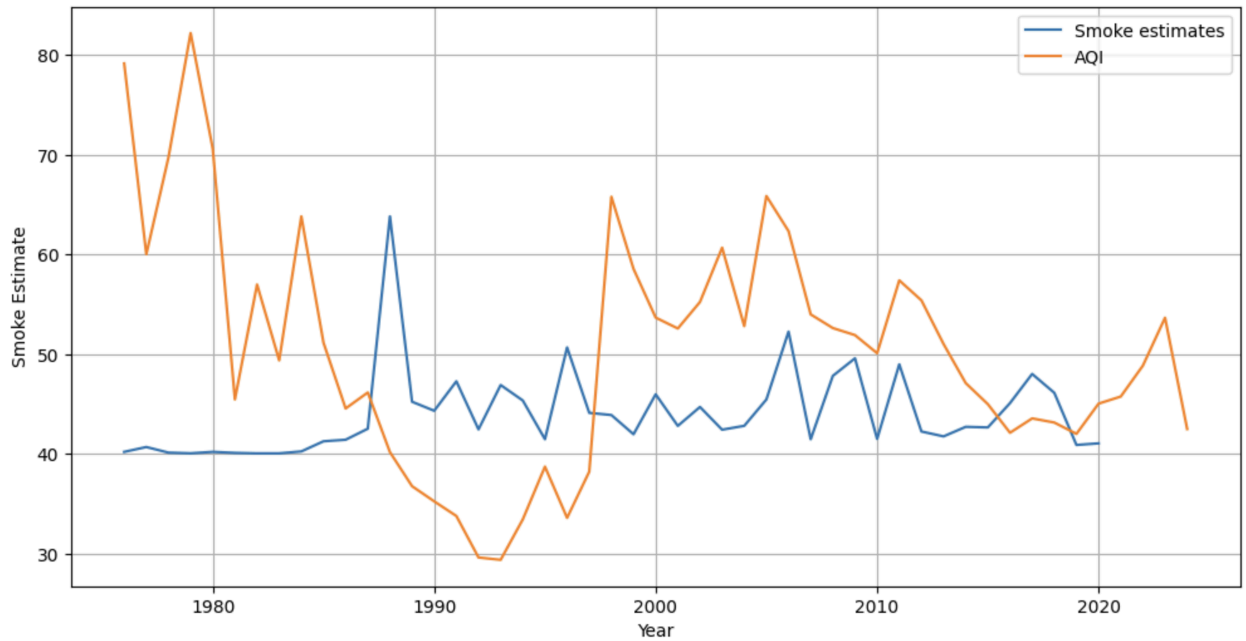

Figure 3: Time series graph containing my fire smoke estimates and the AQI values, for the city of Arlington, Texas.

This plot shows the time series graph for the smoke estimates and Air Quality Index (AQI) values for the city of Arlington, Texas from 1980 to 2020. This data visualization allows for analysis of the trends and patterns in air quality and smoke-related indicators over the past four decades.

We can notice that there are several distinct peaks in both the smoke estimates and AQI values throughout the period. These peaks likely correspond to significant fire events or periods of increased air pollution in the region. For example, the graph shows a particularly high spike in smoke estimates and AQI around the year 1997, suggesting that there was a major fire or air quality incident during that time. Understanding the potential causes and impacts of these peaks could provide valuable insights for fire management, air quality monitoring, and public health efforts in the city.

Another key feature of the graph is the general trend of the data over time. While there are fluctuations and the pattern seems to make no sense at certain points, we can see that at certain peaks, both the AQI and smoke estimates peak together, even if not at the same intensity. This might be indicative of a general correlation between the wildfires and the air quality in Arlington, which is explored further in our analysis. Analyzing multiple similar factors, such as changes in land use, industrial activity, or climate patterns, could help us derive at a much more accurate understanding of how the AQI values worsen as each parameter changes.

Reflection

Working on this project, I learned a lot about wildfires and AQI values, and how smoke might not be the only factor affecting the air quality in a region. Though I was able to see significant trends for certain years, I realized that the smoke estimates calculated using the wildfire data might not always correlate with AQI, which was backed by the conversations that I had with my classmates regarding their results with various cities.

This project required significant collaboration to complete this project's deliverable 1 on time. Our cohort's discord channel served as an important place for communication, where everyone shared their challenges and suggestions for possible solutions at each step. This open exchange of information highlighted common issues that arose, such as the wildfire years not aligning and reported years differing from the actual wildfire years. Hearing about these errors beforehand allowed me to focus more on the analysis itself rather than getting caught up on minor bugs and technicalities.

For the modeling component, I sought guidance from Navya Eedula, who had recently completed an ML internship at a startup and gained expertise in optimization techniques for ML models. Through discussions with Navya, I realized that ARIMAX would be the most suitable model for this problem and not a simple model like multilinear regression. Together, we brainstormed and refined the problem statement and goals, which helped me write the code to get the most optimal solution for my city's estimates. I also collaborated with Abhinav Duvvuri, who helped me overcome the syntax errors and various platform-specific issues that I ran across while working on this project.

Collaborating on this project also enhanced my ability to navigate technical challenges effectively. Undergoing challenges together with others fostered a sense of camaraderie and collective problem-solving. This experience also reinforced the importance of open communication and seeking help when needed. Working with Navya and Abhinav, who had expertise in different areas, exposed me to new perspectives and approaches to problem-solving. This not only helped me complete the project successfully on time, but also expanded my knowledge and skills in data analysis and modeling.