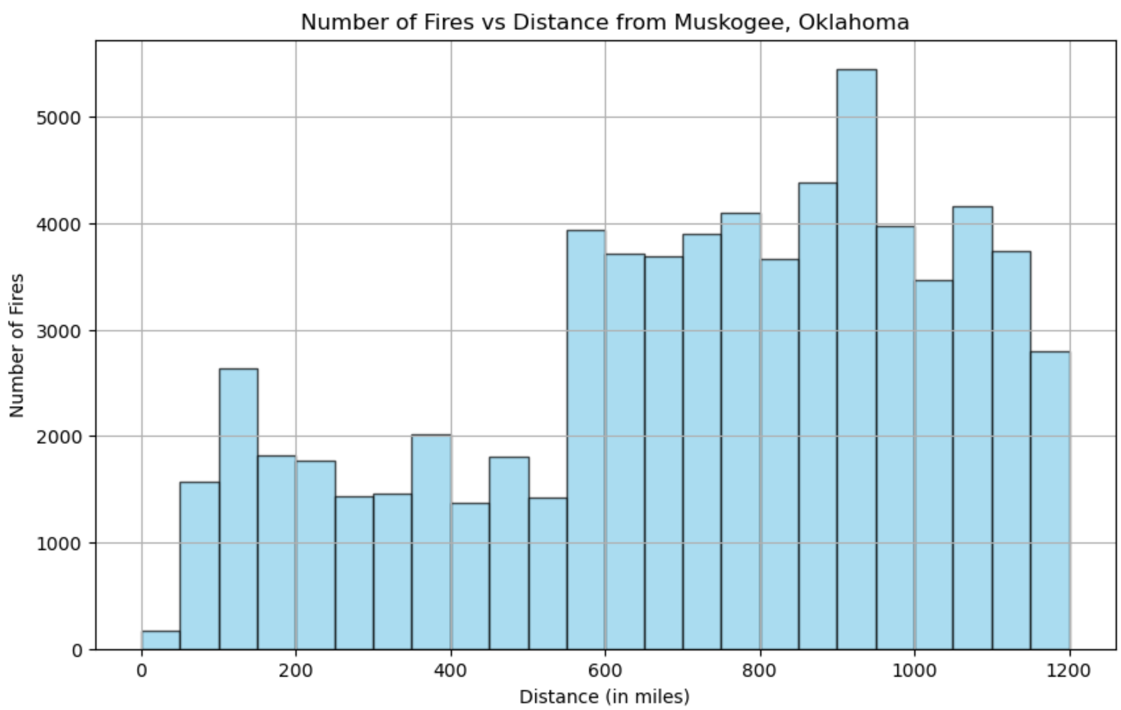
DATA 512 PROJECT

Part 1 - Common Analysis

Visualizations

Histogram for the Number of Fires occurring every 50 mile distance

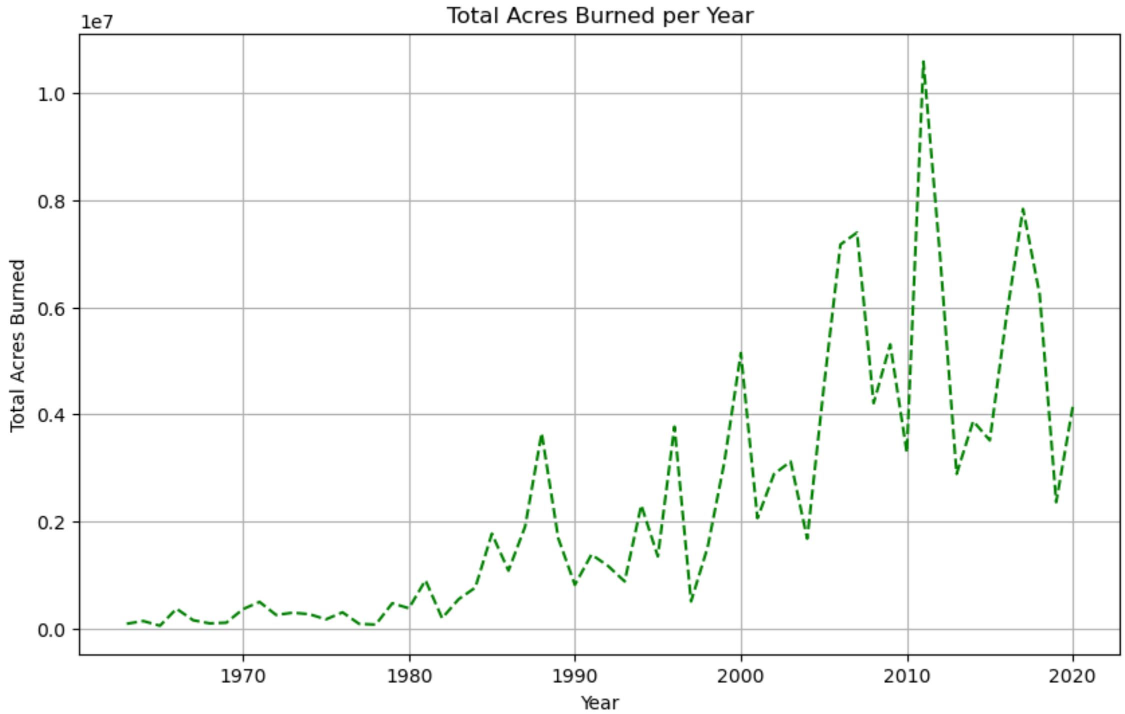


The histogram offers a visual representation of wildfire occurrences within each 50-mile zone, originating from Muskogee and extending up to a maximum distance of 1250 miles. When analyzing the graph, the natural flow is from left to right along the x-axis, which denotes distances in miles. Each vertical bar in the histogram represents a 50-mile interval away from Muskogee, with the height of each bar indicating the frequency of fires within that specific distance range. A taller bar signifies a higher number of fires, while a shorter bar indicates fewer incidents in that particular area.

The x-axis is labeled "Distance (in miles)," providing a spatial context of how far these locations are from Muskogee. Simultaneously, the y-axis, labeled "Number of Fires," quantifies the count of reported fires within each 50-mile distance zone.

Distinct from the typical bell-shaped distribution found in many datasets, this histogram reveals a unique pattern. Specifically, it exhibits a peak around the 900-mile mark from Muskogee. Notably, the majority of reported fires are concentrated within the range of 750 to 1100 miles from Muskogee. This concentration in a specific distance zone underscores a noteworthy spatial trend in wildfire occurrences, providing valuable insights into the geographical distribution of these incidents around Muskogee.

Time Series Graph of Total Acres Burned Per Year

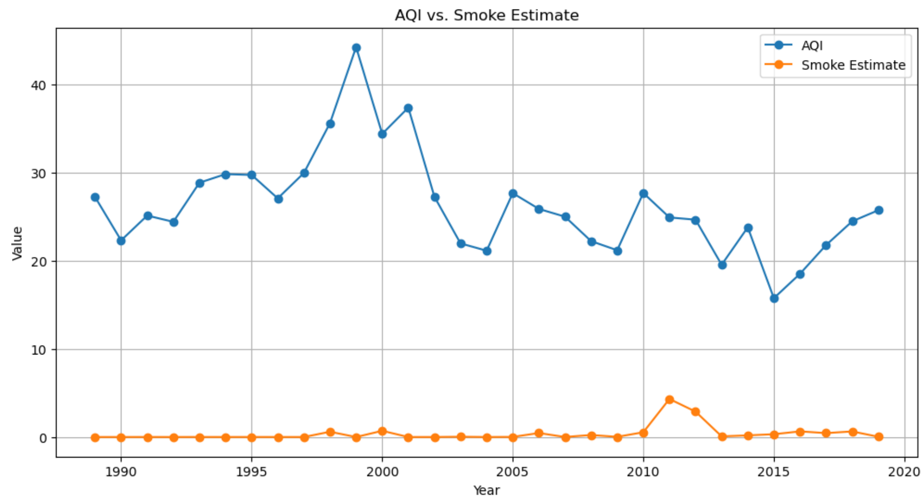


The graph illustrates the annual total acreage consumed by fires within a designated distance from Muskogee, with the maximum limit set at 1250 miles. To comprehend the graph, you navigate the horizontal timeline on the x-axis, representing the years, progressing from left to right. Each data point on the graph corresponds to a specific year, and the vertical axis measures the cumulative acreage burned. As you examine the graph, you can observe how the total burned acreage changes from one year to the next within the specified distance range.
The horizontal axis denotes "Year," marking the passage of time. Simultaneously, the vertical axis quantifies "Total Acres Burned," indicating the combined area affected by fires within the defined bounds.

Upon analyzing the graph, we can see that the total acreage burned as increased over the period of 60 years. In 2011 the graph reveals a peak signaling that 2011 was a particularly challenging year when wildfires caused substantial land damage in the vicinity of Muskogee.

The underlying data for this graph combines Monthly Air Quality Index (AQI) estimates from the US Environmental Protection Agency (US EPA) and wildfire impact data from the USGS. The wildfire data focuses on fires within a 1250-mile radius of Muskogee after 1963. The total acres burned per year were then aggregated, calculated as the sum of burned acres for each year. The resulting dataset is employed to generate the time series graph, offering a visual representation of the yearly variation in acres burned within the specified proximity to the city.

AQI vs Smoke Impact Estimator

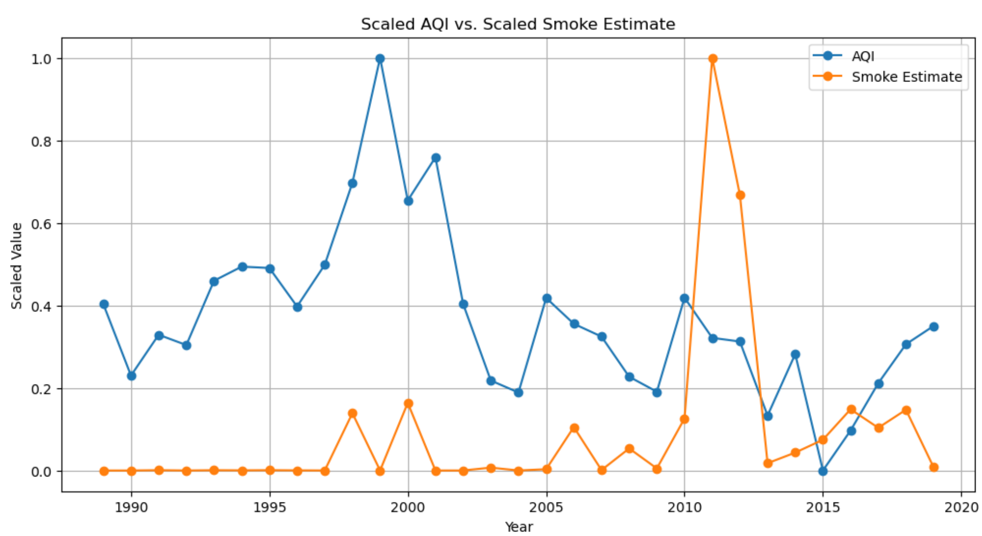


In this visual representation, we present a time series graph depicting two significant estimates — the "Fire Smoke Estimate" and the "AQI (Air Quality Index) Estimate" for Muskogee. This graph facilitates the observation of how these estimates evolve over time.

Direct your attention to the horizontal x-axis, illustrating the years from left to right, offering a chronological timeline. The vertical y-axis measures the respective "Value" estimates. Two distinct lines on the graph, each labeled with a legend, "Smoke Estimate" and "AQI," allow for a comparative analysis of the trends and fluctuations in these two estimates across the years.

Upon initial inspection, a conspicuous positive correlation between these estimates may not be immediately evident. However, further analysis is warranted to thoroughly explore any potential relationships between the Fire Smoke and AQI Estimates for Muskogee.

It's noteworthy that in the subsequent graph, we employ min-max scaling to enhance readability. This scaling technique ensures that both estimates are presented on a standardized scale, facilitating a clearer interpretation of their patterns and allowing for a more nuanced analysis of their interplay over time.



Saumya Nauni

Data Used for the Plots

The dataset utilized in this analysis is a synthesis of information derived from two distinct sources, combining to provide a comprehensive understanding of the environmental conditions in Muskogee:

1. **Monthly Air Quality Index (AQI) Estimates:**
   Sourced directly from the US Environmental Protection Agency (US EPA), this dataset furnishes monthly AQI estimates for Muskogee. Each record in the dataset is characterized by the year and the corresponding AQI values. This information offers a detailed year-wise account of the air quality in Muskogee, serving as a crucial component for assessing the impact of various factors, including wildfires, on the local atmospheric conditions.

2. **Wildfire Smoke Impact Data:**
   Derived from an extensive dataset compiled by the US Geological Survey (USGS), this dataset integrates information on the estimated impact of smoke from wildfires in proximity to Muskogee. The dataset includes key details such as the number of acres burned, the closest distance from the city, and a smoke impact estimate. Notably, the USGS dataset aggregates wildfire information from approximately 40 diverse sources, ensuring a comprehensive and multi-faceted perspective on the occurrences. The focus of this dataset is specifically on the 1963-2023 period and within a radius of 1250 miles from Muskogee, providing a localized and relevant context for assessing the impact of wildfires on air quality.

By merging these datasets, we create a robust foundation for analyzing the interplay between air quality dynamics and the influence of wildfires, allowing for a more nuanced exploration of the factors contributing to environmental conditions in Salina over time.

Reflections

Collaborating on this assignment has been an insightful experience, shedding light on the multifaceted nature of addressing a complex research question. One key takeaway from this collaborative effort is the combination of diverse perspectives allowed for a more holistic exploration of the research question, fostering a deeper understanding of the factors influencing air quality in the context of wildfires.

The use of the *'Wildland Fire Combined Dataset by USGS.json'*, attributed to the US Geological Survey, served as a foundational element for the wildfire impact assessment. Additionally, the incorporation of Monthly Air Quality Index (AQI) estimates from the US Environmental Protection Agency (US EPA) enriched the analysis by providing a comprehensive view of air quality dynamics.

Dr. David McDonald provided valuable code snippets that served as a cornerstone for obtaining both the wildfire and Air Quality Index (AQI) data. Collaborating with peers, including Harshita, Shweta, and Aviva, was instrumental in gaining a deeper understanding of the intricacies involved in processing and interpreting this data. The collaborative discussions extended beyond code implementation, encompassing comprehensive conversations about data processing methodologies.

The collaboration with peers not only aided in technical aspects but also provided an invaluable opportunity to see new perspectives and innovative approaches to problem-solving. By leveraging the collective insights of the team, we were able to navigate challenges more effectively and develop a more comprehensive and nuanced analysis of the interplay between wildfires and air quality.

Throughout the predictive modeling phase, I had the opportunity to leverage a tool called Prophet, which I initially became acquainted with during my summer internship in a business setting. The experience of reusing this tool in an academic environment, distinct from its original application in the business context, proved to be highly enriching. Having previously employed Prophet in a professional setting, adapting its application to an academic research project brought about a valuable perspective shift.

In conclusion, the collaborative activities in this assignment underscored the significance of interdisciplinary collaboration in tackling complex research questions.This experience reinforced the value of collaboration in research endeavors, emphasizing the importance of effective communication and leveraging diverse expertise to address multifaceted challenges.