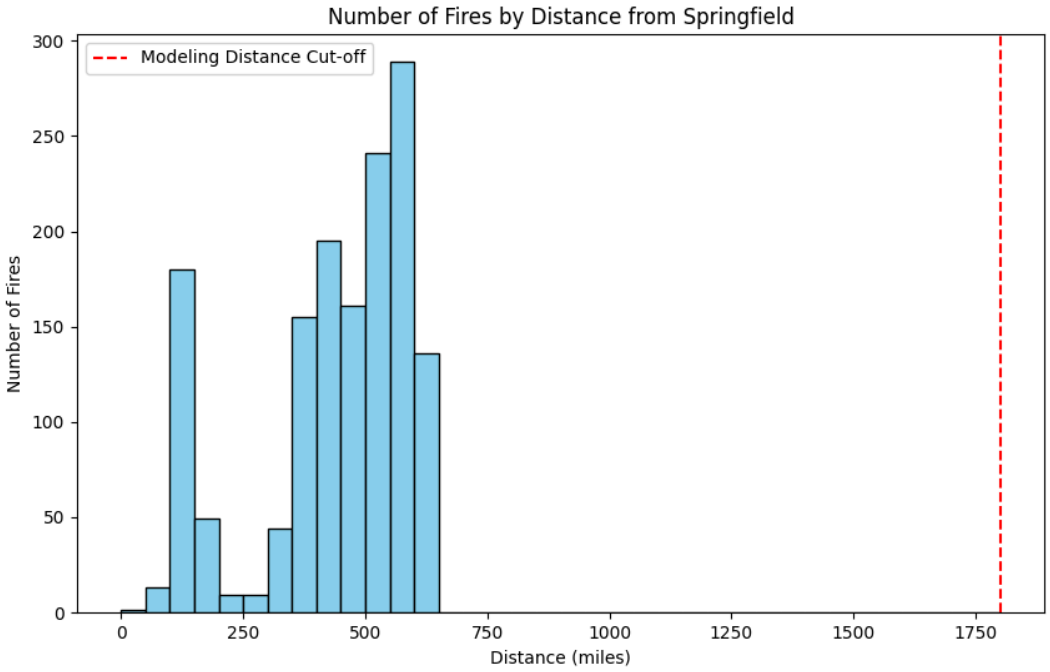
Data 512 Part 1 - Common Analysis

Step 3: Write and reflect

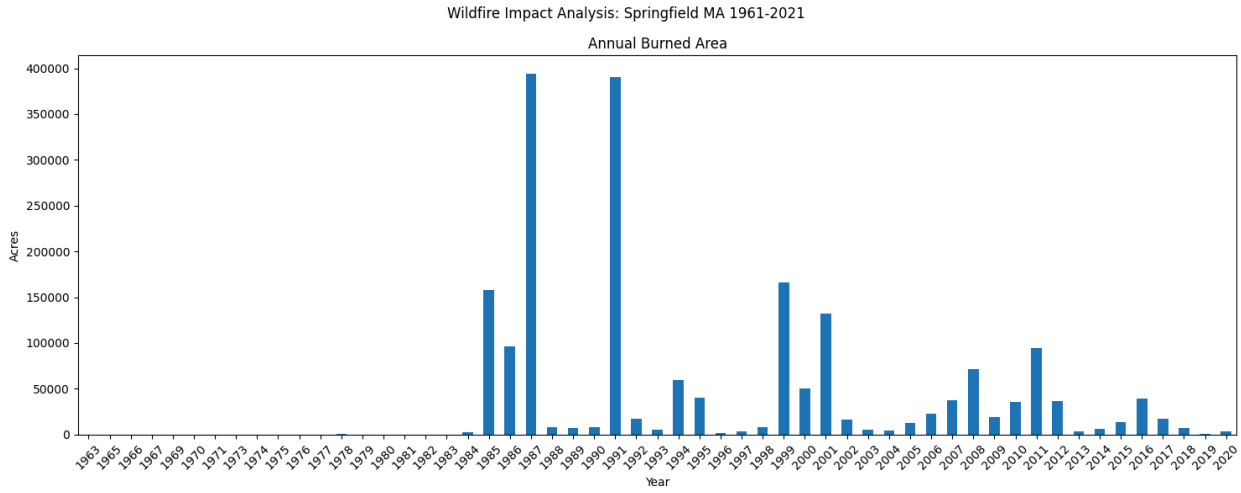
Task 1: Number of Fires by Distance from Springfield



This histogram provides a comprehensive visualization of wildfire spatial distribution in relation to Springfield, Massachusetts from 1961-2021. The x-axis represents distance bands in 50-mile increments extending to 1800 miles from Springfield, while the y-axis displays the count of fires within each distance band. A critical analytical boundary is marked by the vertical red dashed line at 1800 miles, representing the modeling distance cut-off specified in the project requirements.

The distribution reveals several significant patterns and insights. Most notably, there is a pronounced peak in fire frequency between 500-600 miles from Springfield, with approximately 280 fires recorded in this range. The distribution exhibits bimodular characteristics, with a secondary peak around 100-150 miles from the city. This pattern likely reflects two distinct fire-prone geographical regions: one potentially in the southeastern United States (closer peak) and another possibly in the Midwest or southern Appalachian region (farther peak). The relatively low frequency of fires within the first 100 miles of Springfield might be attributed to the predominantly urban and deciduous forest landscape of the northeastern United States, where large wildfires are less common.

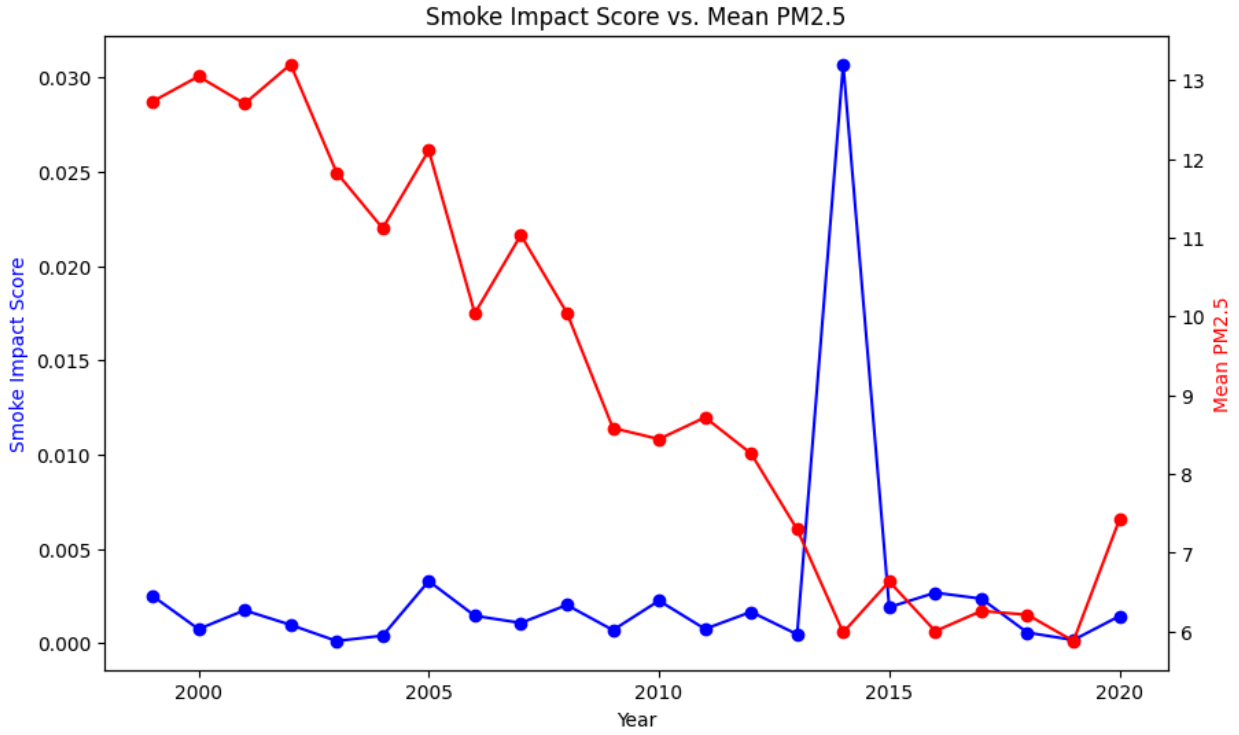
Task 2: Annual Burned Area (1961-2021)



The histogram is an effective way to visualize wildfire distribution concerning Springfield, Massachusetts from 1961-2021. The distance bands extend in 50-mile increments up to 1800 miles from Springfield along the x-axis. In each band, the histogram displays the count of fires that occurred. An important point of significance-a red vertical dashed line marks the 1800 miles-analysis boundary that indicates the modeling distance cutoff specified in the project requirements.

The distribution has led to some audiences being able to give patterns and a few other insights. The largest of these is that a very clear peak in fires occurrence was existent at distances between 500 to 600 miles from Springfield, with approximately 280 fires within this range. The distribution is characterized by two peaks, with the secondary peak being somewhere between 100 and 150 miles from the city. This might indicate two distinct fire-prone geographical areas, one of which could be situated in the southeastern United States (closer peak) and the other possibly in the Midwest or southern Appalachian region (farther peak). The relatively low frequency of fires in the first 100 miles of Springfield may be due to primarily urban and deciduous forest landscapes of the northeastern U.S. where large wildfires are less common.

Figure 3: Smoke Impact Score vs. Mean PM2.5



Dual-axis time series allows us to compare our calculated Smoke Impact Score (blue line, left axis), which is based on measured concentrations of PM2.5 (red line, right axis), taken between 2000 and 2020 for Springfield. Indeed, this graph can let us understand how closely correlated our risk assessment for smoke follows the actual measurements of air quality.

The key feature is the pronounced spike of Smoke Impact Score around 2015, which does not correspond to a similarly prominent spike in PM2.5. PM2.5, indeed, shows a generally declining trend, moving from about 12-13 μg/m³ in 2000 all the way down below7-8μg/m³ by 2020. This divergence would give room for suspicion that the instantaneous assessment method for smoke is coming up with values based on certain other parameters not revealed by PM2.5measured values whereas an improvement in local air quality such as maybe regulated value has captured the impacts of distant wildfires.

Reflection Statement

This collaborative analysis project provided valuable insights into both technical approaches and methodological considerations for environmental data science. One significant learning outcome was understanding the complexities of creating meaningful smoke impact estimates from wildfire data. Through discussions with classmates, I learned various approaches to weighing distance and burn area in impact calculations, which helped refine my own methodology.

The collaborative aspect was particularly helpful in troubleshooting AQI calculations. For instance, while calculating the AQI values for Springfield, I was able to obtain scores, I was able to obtain scores from only 1999. On discussions with my peers, I realized that it was due to lack of sensors integrations across regions and many faced similar challenges.

While collaboration accelerated certain aspects of the analysis, it also highlighted the importance of understanding each component thoroughly.. This approach helped me develop a more nuanced understanding of both the technical implementation and the environmental implications of our analysis.