Analysis Reflection

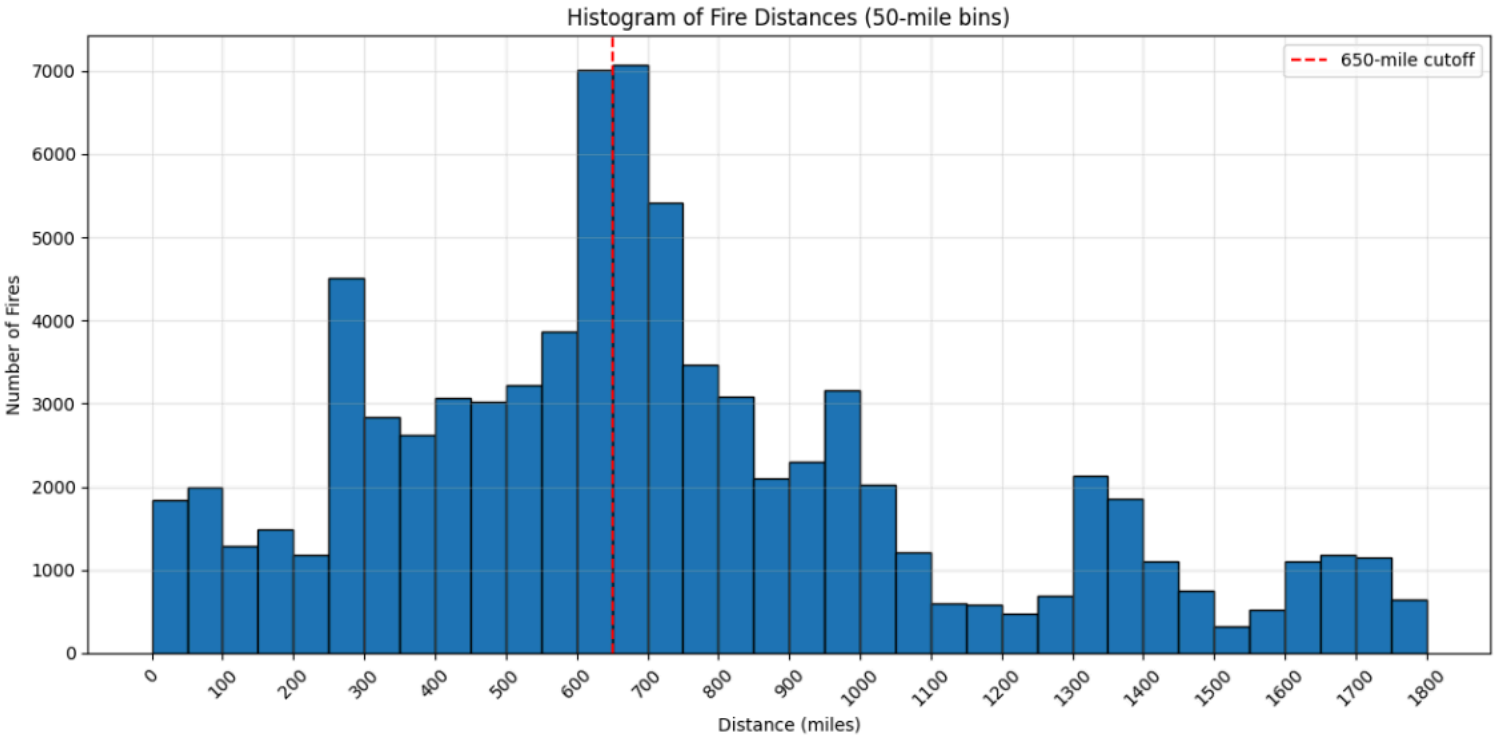
Chakita Muttaraju

10/20/2024

Reflection on data visualizations

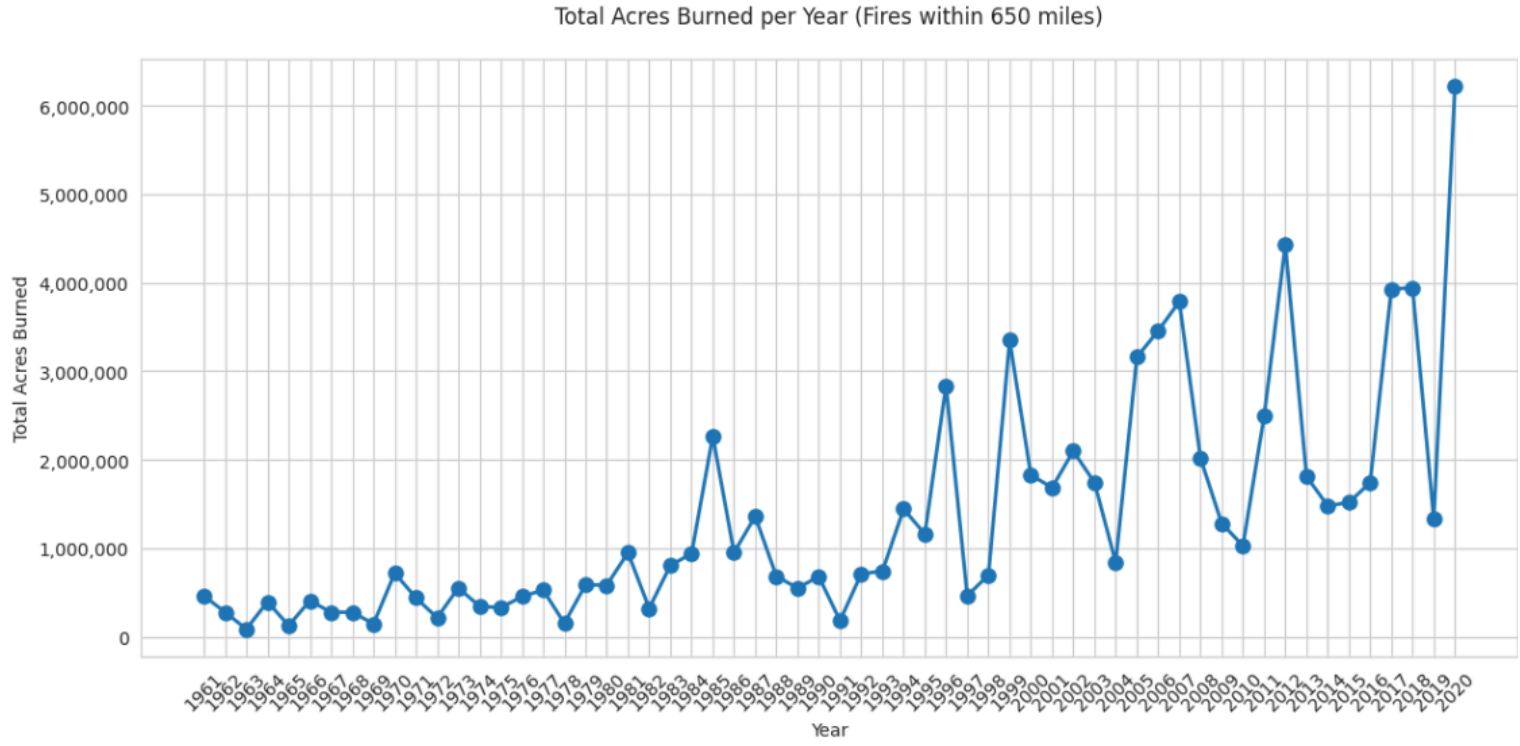
1. **Histogram showing the number of fires occurring every 50 mile distance from Lubbock, Texas for all fires ranging up to 1800 miles**

   The histogram represents the cumulative number of fires occurring at every 50 mile interval from Lubbock, Texas up to 1800 miles. The x-axis represents the distance of the fire from Lubbock in miles and the y-axis represents the total number of fires at that distance. The distance of the fire from Lubbock was calculated from geographic coordinates of the fire provided in the USGS data for Combined wildland fire datasets for the United States and certain territories, 1800s-Present. Since the earth is ellipsoid, the distance between geographic points is more of an arc rather than straight line distance so we use PyProj, a python module to transform our distances to get a straight line distance between the fires and Lubbock. The red vertical (at 650 miles) line represents the cut-off that we used in calculating the smoke estimate



- a metric to represent the impact of wildfires on Lubbock. From the histogram, we can see that 650 miles is around the mean value and most of our data lies within 650 or 700 miles away from Lubbock.

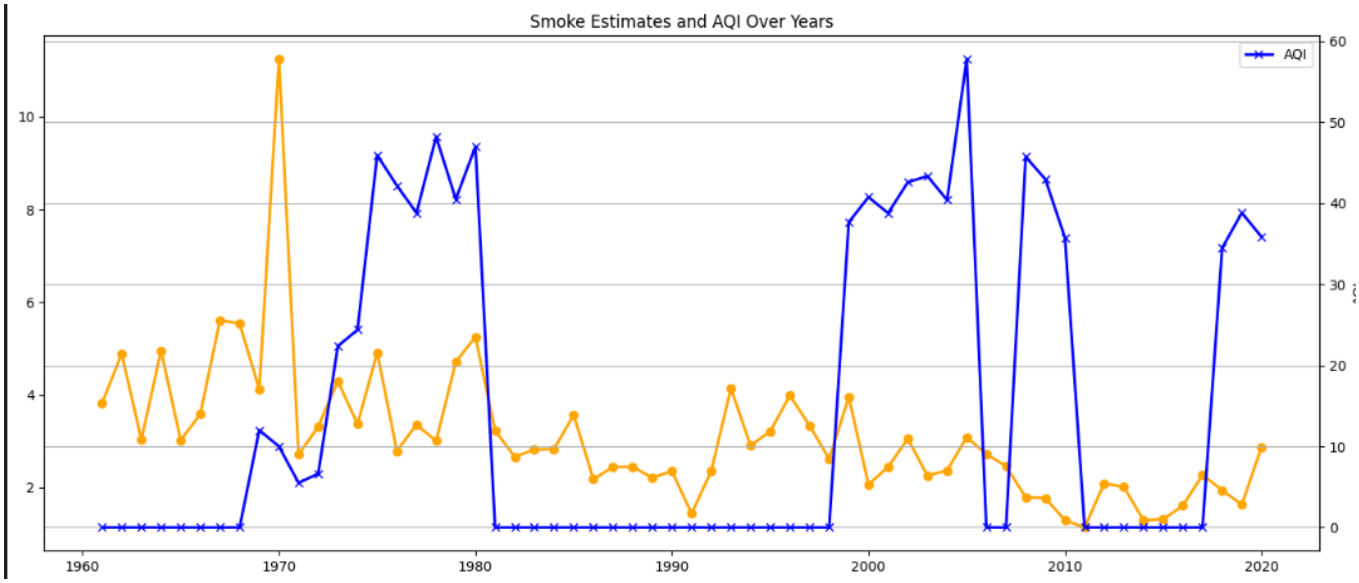
2. **Time series graph of total acres burned per year for the fires occurring in the specified distance from Lubbock, Texas.**
   The graph below shows the trends in the acres of land burnt over the years ranging from 1961 t0 2020. The x-axis shows the different years and the y-axis represents the total GIS acres of land burnt in wildfires. The [USGS data for Combined wildland fire datasets for the United States and certain territories, 1800s-Present] data contains data from 1800 to 2020 but we limit our analysis to the last 60 years which is from 1961 to 2020. From the plot, we can observe that the total acres of land burned per year due to wildfires has steadily risen with a particularly sharp rise from 2019 to 2020. The total acres of land burned or the size of the wildfire is the second component in our smoke estimate after the distance of the wildfire from Lubbock. We



hypothesize that the size of the wildfire is directly proportional to its impact on the cities around it.

3. Time series graph containing fire smoke estimates and the AQI estimates for Lubbock, Texas

The graph below shows the time series trends of AQI estimates and the smoke estimates for Lubbock from 1961 to 2020. The smoke estimates were calculated using the distance of the fire from the city as well as the size and type of wildfire. The AQI data was pulled from the US EPA Air Quality System (AQS) API. The x-axis of the graph represents the year and we have twin y-axes for smoke estimate on the left and AQI on the right since our smoke estimate and AQI are on different scales. The points on the AQI graph where it falls flat, the AQI data was unavailable for that year and was hence considered to be zero. We can observe that the AQI and smoke estimates both trend downward indicating that smoke impact on Lubbock has been decreasing over the years. However we observe from the total acres graph that the total land burned has risen steadily so this could just be a variation in the data and might rise again in the future or it could also potentially indicate that the fires have shifted a bit farther away from Lubbock over the years.



Reflection on collaboration aspect of the project

I think this project involved a lot of collaboration from Prof.David as well as from my peers. Firstly, the example code notebooks provided by Dr.David McDonald were extremely useful especially since I don't have much expertise in this area of geographic data. One important aspect of collaboration for this project was building upon his code to extract the required data and run the analysis. Secondly, since the project is fairly open-ended, I collaborated with my peers in discussing potential methods to calculate smoke estimates and statistical models for time series analysis to predict smoke estimates in the future. Since the GeoJSON data required a lot of processing, I also collaborated with my peers and

discussed strategies to extract start dates and end dates for fires in order to determine if they occur during the fire season.