STA 160 Final Project Reports

Niraj Bangari, Jayoung Kim, Melanie Lue, Cecilia Nguyen, and Marvin Pepito

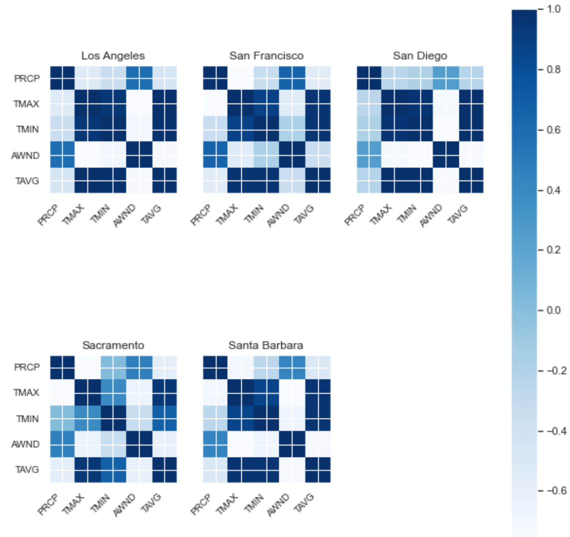
I. INTRODUCTION / BACKGROUND

A. Background

The Global Summary of the Year (GSOY) is a dataset containing climate data from various locations in the world. There are 58 climatological variables, as well as thousands of locations. These locations are named with their city or airport name. The climatological variables are computed from the summary of the day results in the Global Historical Climatology Network Daily (GHCN-D) dataset.

B. Research question

Our research question is that we want to analyze which general areas of California are more fire-prone over a 20 -year period . We would look at the factors that cause wildfires such as high temperature, low precipitation, and increased wind speeds. Our variables chosen would be: Average Annual Precipitation, in inches (PRCP), Average Annual Temperature, in Fahrenheit (TAVG), Average Maximum Temperature, taken as an average of mean monthly maximum temperatures in Fahrenheit (TMAX), Average Minimum Temperature, taken as an average of mean monthly minimum temperatures in Fahrenheit (TMIN), and Annual Average Wind Speed, in miles per hour (AWND).



C. Correlation Analysis

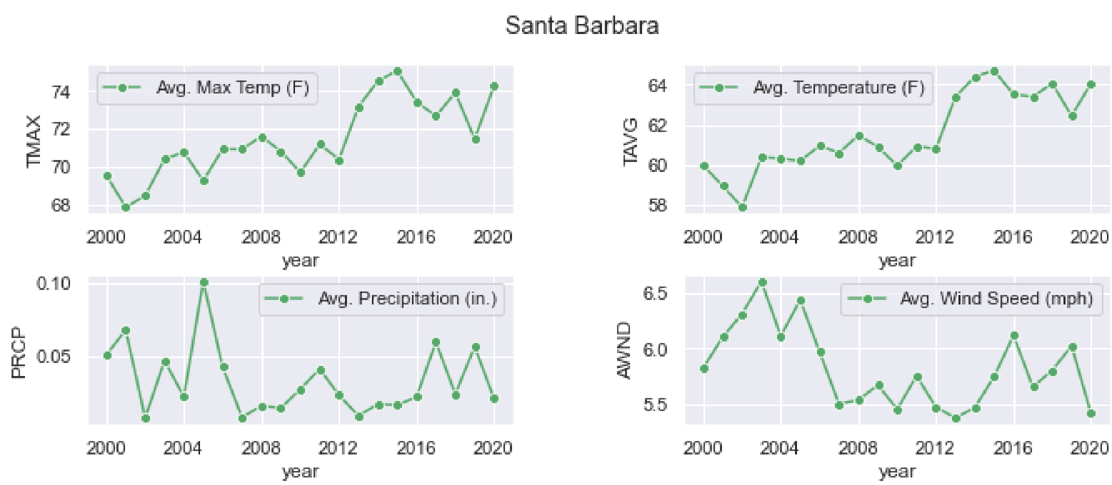
To explore the relationships between variables and the strength of the correlation between them, we analyzed each city. The correlation strengths in Sacramento are slightly different from the other cities; this might be due to the fact Sacramento is a bit more inland. For most cities, we can observe moderate correlations of ranging around .2 to .6 between the variables that we are interested in: Average Max Temperature and Weighted Average Wind Speed, Average Temperature and Weighted Average Wind Speed, Precipitation and Average Max Temperature, and Precipitation and Average Temperature. We can see a moderate negative correlation between both Average Max Temperature and Weighted Average Wind Speed (where the higher the temperature the lower the wind speed), and also a moderate negative correlation between Precipitation and Average Max Temperature, and Precipitation and Average Temperature (where the higher the temperature there is less precipitation).

D. Why did we choose these cities?

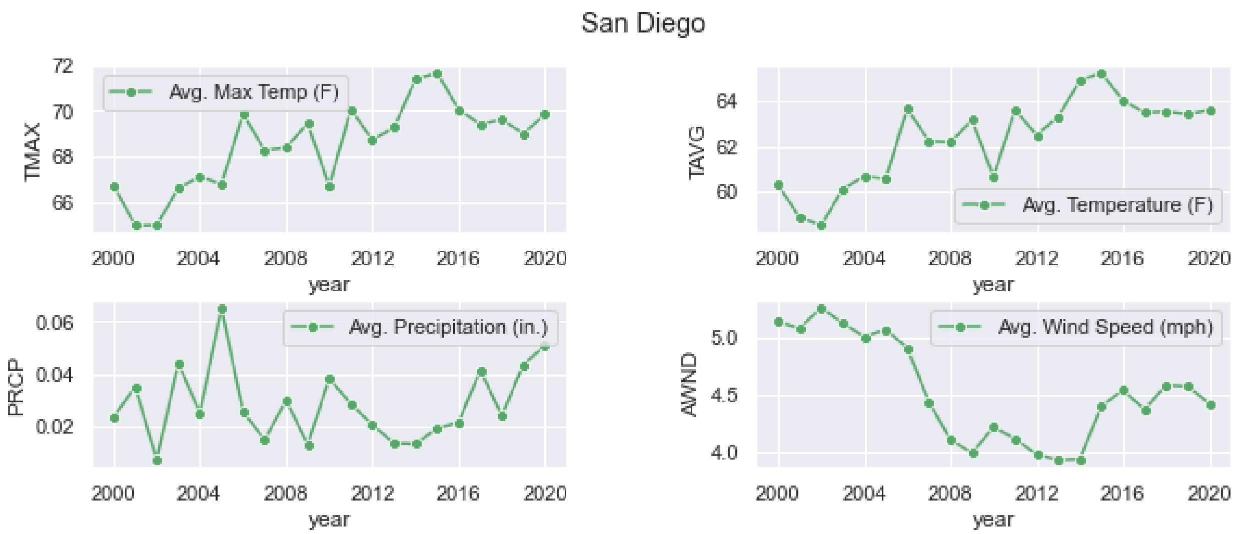
We chose stations throughout California in order to gain information about the general region (Northern California, Bay Area, Southern California). We didn't want to choose cities that were extremely near each other in order to analyze cities with different weather conditions. To check if our interpretations of our plot for the trends of these variables over a 20-year period are sensible for each city and its general location, we will compare our data with a chosen year (2007) with a geospatial visualization of California depicting the density of wildfires (2007).

II. BY CITIES

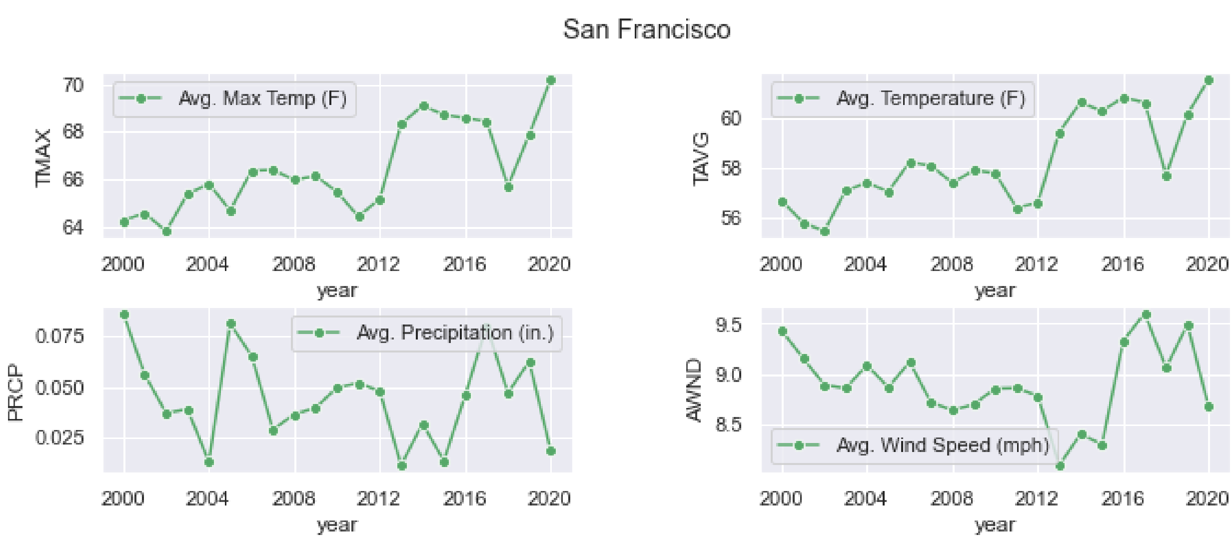
Now we will explore the trends of our four variables for each of our California cities.



For Santa Barbara, we can see that both average maximum temperatures and average overall temperatures during our 20 year period peaked in 2015, with lowest temperatures in 2002. Average precipitation in Santa Barbara reached its highest point in 2005, however this peak is not reached as gradually as that for temperatures, signifying an abnormally “wet” year for 2005. Around this time, we also see a dip in the average temperature, which was on an otherwise steady rise from 2003 to 2006. We would expect this since there was much higher precipitation in 2005 than adjacent years. Average wind speed in Santa Barbara is more volatile year over year than our other variables, with multiple sharp peaks over our 20 year period.

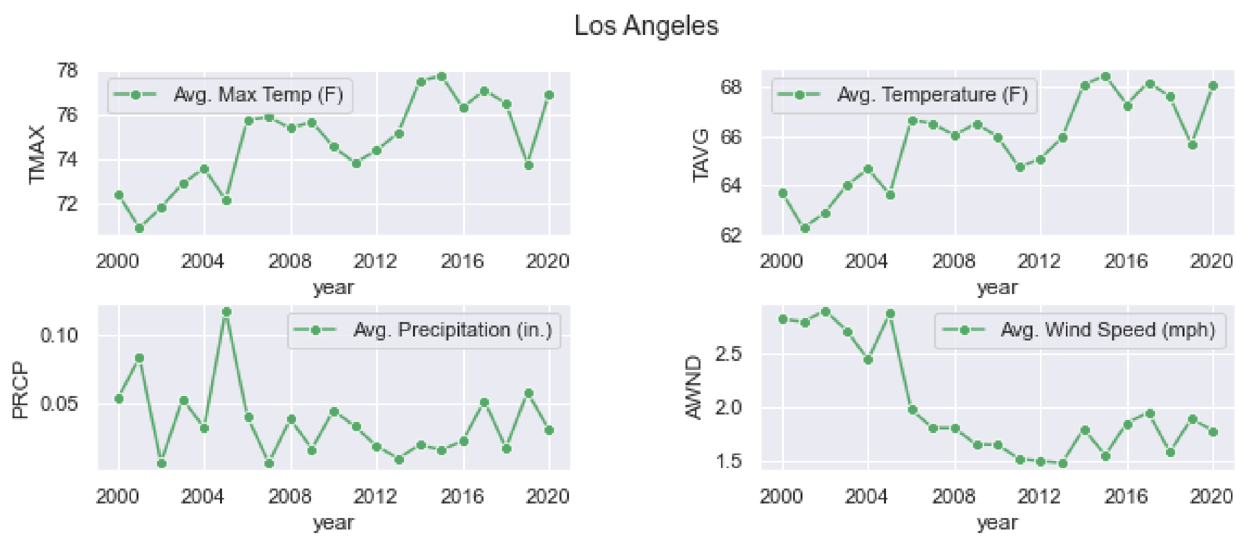


San Diego has a similar peak in temperatures around 2015. However, it is a few degrees cooler than Santa Barbara, on average, in terms of maximum temperature. We also see a similar pronounced peak in precipitation around 2005, with average wind speeds decreasing steadily after 2005 to about 2009. In San Diego, we see a very drastic dip in temperatures in 2010, corresponding to a relative jump in average precipitation and wind speed for that year.

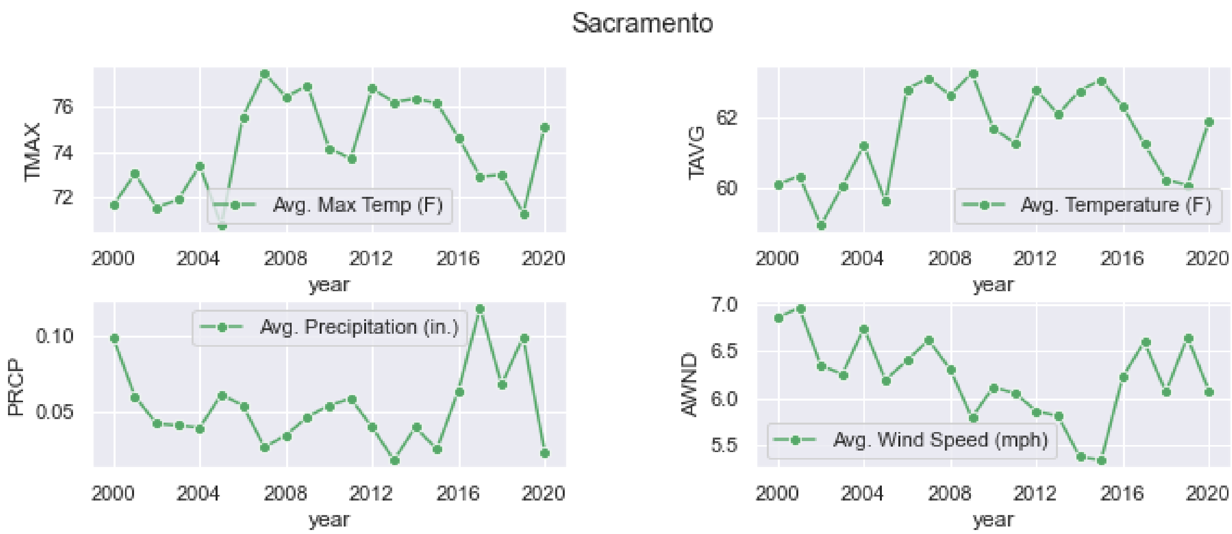


San Francisco displays a similar trend for temperatures in our 20 year period, however it is a much cooler city overall than San Diego and Santa Barbara. For this reason, we see the trend in our plot shifted down in degrees Fahrenheit. One major difference in San Francisco’s temperature trend is the drastic dip in

2018, which was cooler than the previous year by almost 3 degrees, on average. San Francisco also received its peak precipitation in 2005, however with additional heavy precipitation in both 2000 and 2017. Average wind speed is also much higher than the previous two cities across our entire time period, with a peak of over 9.5 mph. We can attribute these observations to San Francisco being more north on the California coast, with different weather patterns coming from the Pacific Ocean than southern California.



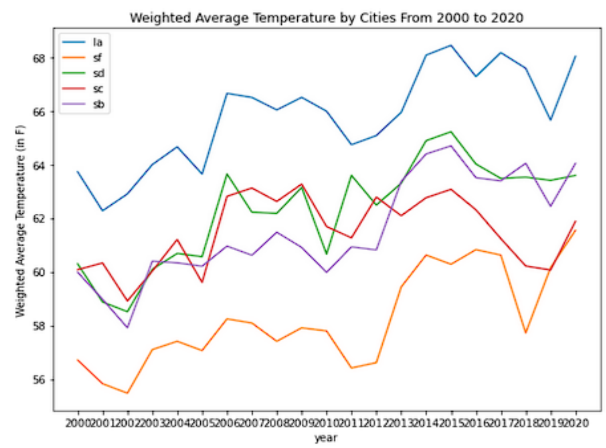
For Los Angeles, over the 20 year period, the average maximum temperature increased steadily from the low 70s range to the high 70s from around the years 2004 to 2020. This may indicate that those years experienced were slightly higher than normal, compared to the years before and after. We also see that average temperatures increase steadily to values around the high 60s. Along with the steady increase in average maximum temperature, we can also observe that the average wind speed decreases steadily to around the values between 1.5 mph and 2 mph near the end of our observed time period. This period also had a steady decrease in precipitation to levels around .05 inches and less.

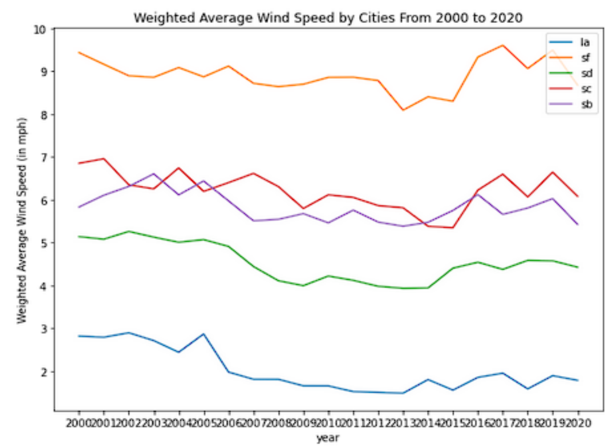


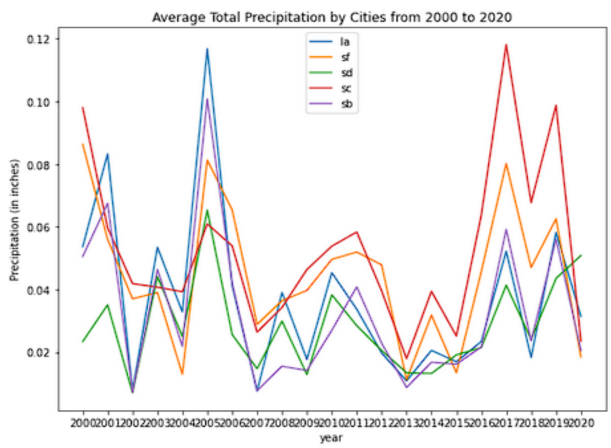
For Sacramento, over the 20 year period, the average maximum temperature increased to about the mid to high 70 degrees Fahrenheit from around the years 2004 to 2020. Average temperature also increased by a few degrees; both average maximum temperature and average temperature is observed as a peak compared to the years before and after. With the increase in temperature, wind also steadily decreases from around 6.5 mph to a little below 5.5 mph. We can observe that during this period, precipitation levels decreased to less than .10 inches.

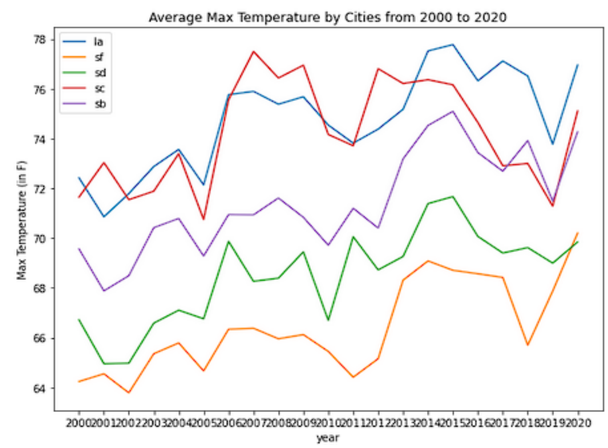
For all the cities, there is a similar trend in the variables for each city where the average precipitation decreased, the average temperature increased, and the Weighted Average Wind Speed decreased. These trends make sense since the hotter the temperature, the less precipitation and wind is present. We see from these plots that San Francisco is the coolest of our five cities. Los Angeles and Sacramento are the two hottest cities in our group, with consistently higher average temperatures and max temperatures than the other cities. The coastal cities have similar trends in precipitation, all with peaks around 2005, while Sacramento, the only inland valley city in our consideration, has much lower precipitation overall and does not follow a similar pattern. Additionally, average wind speed in the northern california cities seem to be higher in volatility and magnitude than the three southern california cities.

III. BY VARIABLE



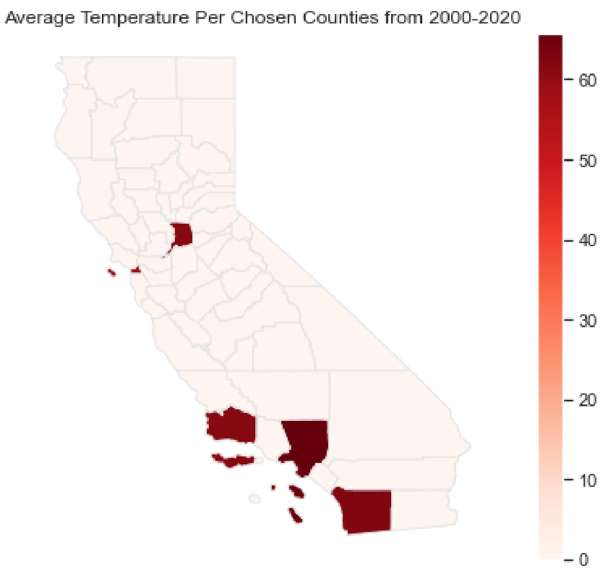


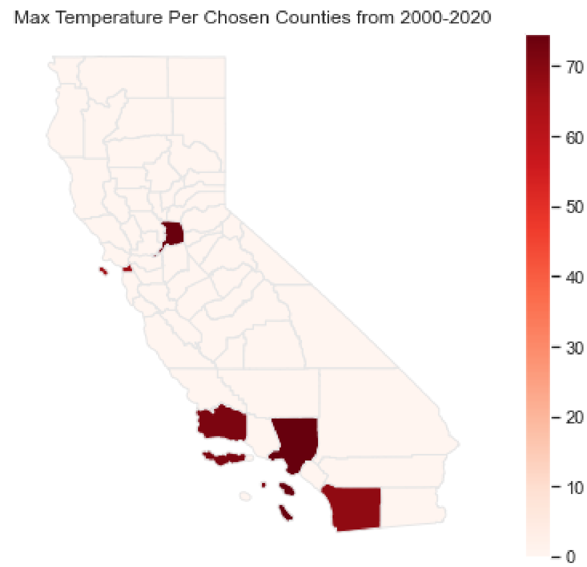


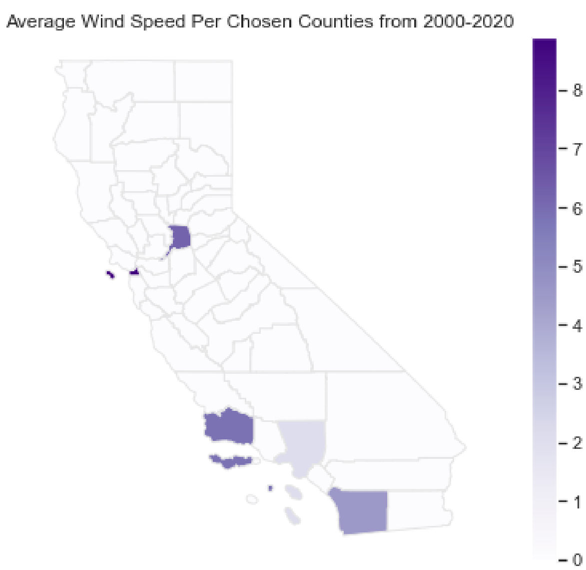


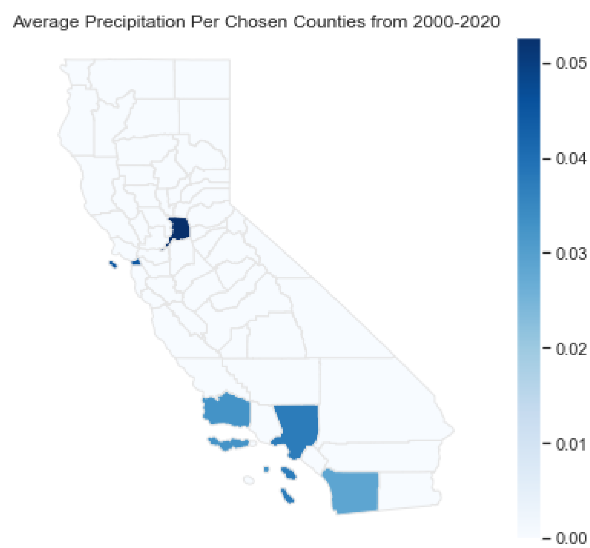
For each variable, each city is plotted over each other so that we can compare the cities trends against one another. In the plots of all of the variables, weighted average temperature, weighted average wind speed, average total precipitation, and average max temperature, the five cities all have very similar trends. In the weighted average temperature plot, all cities, with the exception of San Diego, have dips around the years 2002, 2005, 2011, 2019. San Diego's temperatures seem to be more sporadic, but in general, all of the cities' trends are generally positive. Los Angeles has the highest weighted average temperature, and San Francisco has the lowest. In the weighted average wind speed plot, the cities all generally decrease until around 2015, and then increase. San Francisco and Sacramento have more of an increase, and unlike the other cities, Santa Barbara has a brief positive trend in the beginning. San Francisco has the highest weighted average wind speed, and Los Angeles has the lowest. For the Average Total precipitation plot, Sacramento and San Francisco start with a negative slope, while the rest of the cities are positive, however after around 2002, Santa Barbara had similar trends as Sacramento and San Francisco. After about 2011, all cities had similar trends. Los Angeles and Sacramento have the highest average total precipitation. In the average max temperature plot, the trends are very similar to the weight average temperature plot. Similarly to the weighted average temperature plot, Los Angeles and San Francisco have the highest and lowest temperatures, respectively, but Sacramento also has one of the highest temperatures.

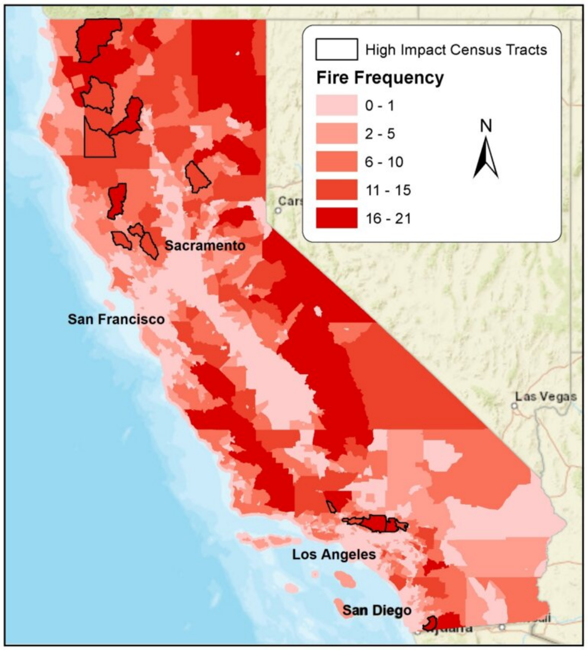
IV. GEOSPATIAL VISUALIZATION











Data from UCI

To see the relationship between the variables such as average temperature, max temperature, average wind speed, and average precipitation with the frequency of the wildfire and compare them geospatially, a shapefile containing the coordinates of California counties is used to visualize each variable into maps and the datasets for each city are merged by each county into the shapefile.

The top four maps below show the average temperature, max temperature, average wind speed, and average precipitation per chosen county from 2000 to 2020 respectively which are considered as the main factors that cause wildfires such as high temperature, low precipitation, and high wind speed. The last map is from the study by University of California, Irvine [2] which analyzes the fire frequency and area burned by wildfires across all Census tracts in California over the last 20 years and visualizes the result into the map.

The average and max temperature maps almost look the same where the counties with high max temperature also have high average temperature as well so it seems considering both maps is not efficient in this case. The temperature maps have a clear pattern in Northern California areas such as San Francisco and Sacramento where they have lower average temperatures as the fire frequency in that area is low. The

Northern California areas also have a noticeable pattern with the wind speed and precipitation maps where the fire frequency is low they have higher wind speed and precipitation then the areas in Southern California such as Los Angeles and San Diego areas.

V. CONCLUSION

In our project, we analyzed which general areas of California are more fire-prone for a 20 year period . We analyzed factors that cause wildfires such as high temperature, low precipitation, and increased wind speeds by extracting temperature, wind, and precipitation values from the NOAA API using the requests module in Python.

After combining the data and interpreting several different visualization methods, such as correlation matrices, multiple line plots, and geospatial plots, we found that areas with a lower fire frequency, such as San Francisco and Sacramento, had higher wind speed and precipitation levels then areas such as San Diego and Los Angeles, as seen in both Section 3 and 4.

These sections also show the unsurprising relationship between average temperature and fire frequency, as once again, San Francisco and Sacramento have the lowest recent values for average temperature, and are both areas with low fire frequency. Interestingly enough, the maximum temperature attribute does not seem to be as effective as a predictor, with Sacramento having the second highest values recently as well as a dark red color on the geospatial plot. This allows us to infer that consistently higher temperatures over time are a more significant factor than high but inconsistent peaks.

VI. REFERENCES / CODES

[1]National Centers for Environmental Information (NCEI). “Global Summary of the Year (GSOY), Version 1.” *NCEI*, https://www.ncei.noaa.gov/metadata/geoportal/rest/metadata/item/gov.noaa.ncdc:C00947/html.

[2]Bell, Brian. “California Wildfires Disproportionately Affect Elderly and Poor Residents, Study Finds.” *Phys.org*, Phys.org, 31 May 2021, https://phys.org/news/2021-05-california-wildfires-disproportionately-affect-elderly.html.