Soham Shah

CSCI 7000 - Scientific Visualization

Tim Dunn

Final Project Report

Work Throughout the Semester:

I got the data ~2 weeks before the final project presentation was due. Also, I was expecting another dataset, but that didn’t work out, so I just used the data I had to create my visualizations. Because of this, my work ended up being very last minute, instead of being spread out over the course of the semester. One of the big problems I ran into was that I really liked my original idea – which was about showing municipal bond interest rates on a map over time and potentially tying it to natural disasters. But I didn’t have the data, and after researching more, I realized that it was very difficult to get access the data. So, I needed to change the goal and scope of the project and ended up switching to fire data.

Workflow Used:

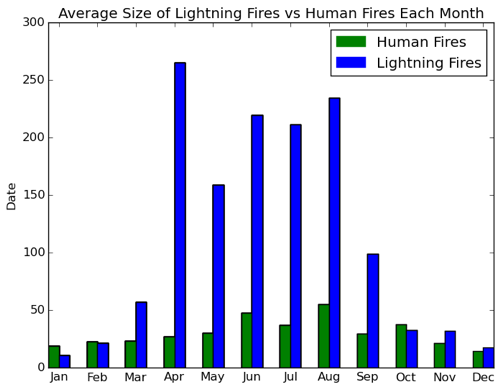
The workflow I used was pretty simple. The data was all in CSV files, so I first figured out how to get the data into pandas using the read\_csv() function. Once I knew I had the data properly reading, I started manipulating it and slicing it using pandas operators before feeding it into matplotlib. I started off by creating basic bar charts and pie charts using matplotlib. If I liked something that I generated then I would start saving those images as plates so that I could add them to the movie. I often worked backward as well. For example, after looking at the data, I decided It would helpful if the viewer knew this about the count of fires for each type. Then, I figured out how to best represent that using a visualization.

For my basemap work, I slowly built up my code to work for the entire dataset. I first worked on plotting a map of the US using basemap. Then, I took a very small sample -5 points – from the dataset and worked on mapping them using latitude and longitude. From there I got the size and shape of the dot to represent the size of the fire. Finally, I transferred the code over to the Supercomputer where I ran all the rendering, since I knew it would take a long time if I ran it locally. In order to get all the plots back to my local computer for post processing, I zipped the whole folder and uploaded it to google drive.

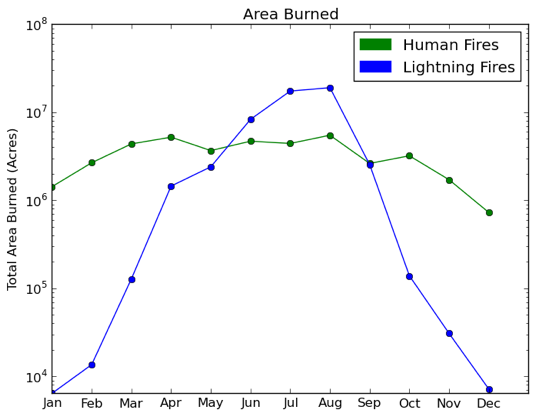
Once all the plots were done, I moved into visualization. I had a lot of issues with getting Natron to work, so I used Adobe Premiere Pro for creating the final movie.

The Story to describe the science:

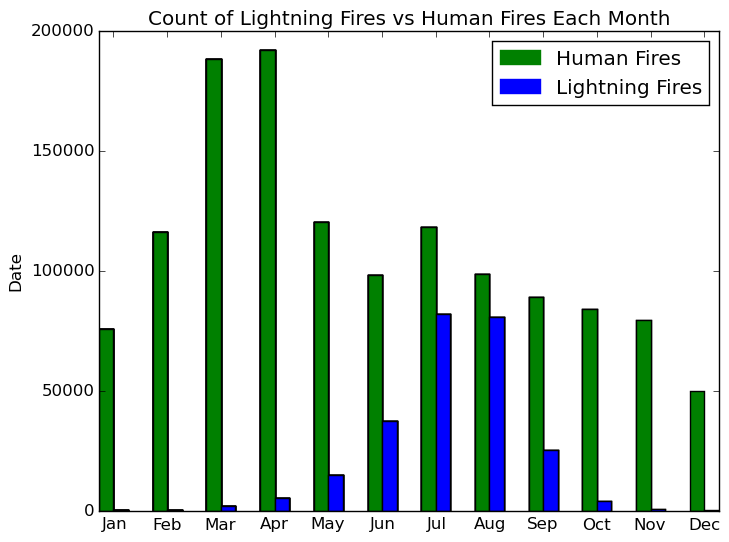
The goal of the project was to help show the difference between human-caused fires and lightning-caused fires. I wanted to show how lightning fires are seasonal and occur very frequently during the summer. It was interesting to look at the size of both types fires as well. I also wanted to show how the location of these fires differs. I let viewers draw their own conclusion about the data. The main goals was for viewers to understand how fires happen in the US. I showed things like the average size of human-caused fires versus lightning caused fires each month:



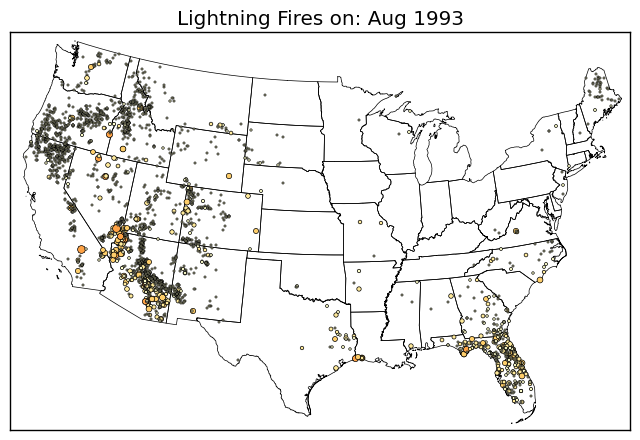
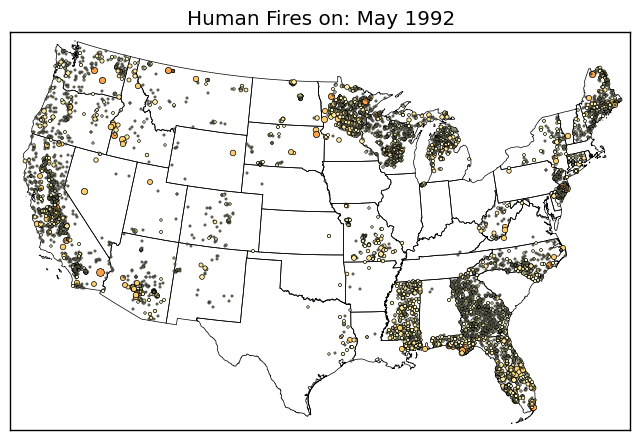
The total area burned each month in acres by each type of fire:



The count of each type of fire



A key highlight of this presentation was showing the plots of where each type of were located on a map of the United States where the size and color of the dot indicate the size of the fire.



What you learned when producing the final presentation:

This was a huge learning opportunity for me. First, I had never worked with pandas before, so it was interesting to figure out how to get data into a dataframe and then use slicing operators to pull out the data I wanted.

While I had worked with matplotlib a little bit, I had not ever used it to generate visualizations. It was an interesting learning process to figure out how to get matplotlib to generate plates that could then be used in the movie. I’m not sure if my method is correct, but I created a for loop where I would keep the first column of data in a row and 0 out the rest so that the plot would still be sized correctly. For the next plot, I’d keep two columns, and so on until it generated the entire plot.

Similarly, for basemap, I had never used it before. I got to learn how to use shape files for generating the map itself. I also had to learn about how to take latitude and longitude and put it in the format basemap wants for plotting. Finally, I learned about how to create custom color tables so that the color reflected the size of the fire.

Finally, I got to learn how to use Premiere Pro. I had never used it before, but I’m familiar with Photoshop, and I’m glad some of that knowledge transferred. It turned out to be lucky because it worked much better than Natron at doing the visualizations.

What you learned in class that helped most

The most helpful thing that I learned in class was the visualization workflow. Before, I had no idea that you had to generate plots for different times in the visualization. Also, I used the outline of the python code we wrote for generating plots in visit and ported that over in order to generate plots using matplotlib and basemap. The concept of having a for loop and a save image function is still the same.

The class was also helpful because it helped build my intuition about how to present the data in a way the viewer can understand. I realized that if I put all the fires onto a map of the US, then it wouldn’t mean anything. There would be too many points and the viewers would not be able to tell what it meant or figure out any trends in the data. Even trying to do it by year seemed too dense, which is why I settled for months.

What you would have liked to improve on, or add to the presentation, given more time and resources:

Given more time, I would have liked to make the movie nicer. Because I spent so long trying to get Natron to work, when I switched over to Premiere Pro, I didn’t spend much time creating transitions that looked nice or good looking title slides. I just went with the default in order to get something that worked in the limited time I had remaining. Given more time, I think I could have gotten some very nice fire images in the presentation to make it look more visually appealing. Also, I would have liked to create some more interesting visualizations. It would have been cool to look at the trend of the number of fires over the years and maybe take a granular look at some very bad fire seasons to see what the effects were. Something else I didn’t get to show was the July 4th effect. Which is where there is a huge uptick of human-caused fires on that day because of fireworks.