Hotshot Crew 554 - US Wildfires Analysis*

Asumi Suguro[1], Kaile Huang[1], Luke Nelson[1], and Po-Han Chen[1]

[1]University of Southern California, Los Angeles CA 90007, USA
suguro@usc.edu, kailehua@usc.edu, lukenels@usc.edu, cpohan@usc.edu

Abstract. This paper describes our application, US Wildfires Analysis, which recommends gearing efforts to fire prevention as opposed to merely suppression through the progression of six visuals and data from 1992-2015.

Keywords: wildfires · fire prevention · fire suppression · wildfire causes · wildfire seasonality.

1 Introduction

Forest fires in the western United States have been increasing in both size and severity for several decades [**?**]. These increases have been attributed to human activities and climate change. With all of the recent fires the US has been experiencing, this information is top of mind for health authorities, fire departments, environmental organizations, and even the common person such as you or me. The goal of our project is to first provide users with an understanding of the recent growth in wildfire incidence and severity. Secondly, we promote action over reaction as it relates to fighting wildfires. Historic patterns and relationships among wildfire causes, seasonality, location, and spending support the notion that being proactive towards preventing fires is achievable. This approach of working to prevent wildfires should be given just as much weight and attention as is given to putting out fires that have already started. Our application drives home this goal via a progression through six visuals. The insights provided in our project will be pertinent to both the private and public sectors. Government agencies and fire authorities, as well as the general public, will be able to utilize the information presented to better identify situational risks and target activities that are more likely to result in sparking a wildfire. There is certainly extant work pertaining to US wildfire growth over the past few decades, however, we have contributed something both useful and unique. While there are a number of US wildfire related dashboards in existence, we haven't found anything as comprehensive as ours that lets you interact with and walk through the data, in addition to telling a story along the way. The other softwares we've seen simply allow the user to play around with the data on their own - however they are often not intuitive and are quite archaic. The contribution of our application is that it walks the user through an interactive story. For example, LA Times pushed a California specific infographic, but it does not provide the ability to explore the data or interact with the visualizations [13]. We believe that our design helps with memory and gets readers engaged with the information and what is going on. We will dive into related work further in this paper. We are not only giving our users data, we are giving them a mission.

In this paper, we will be discussing the intricacies of our application and how it was contrived. in Section 2, we will describe Related Works in terms of how our application differs from them and accommodates for their shortcomings. Section 3 discusses the Data we used and the steps we took to process it. Our overall Approach, which surrounds design and implementation, is discussed in Section 4. Section 5 covers aspects related to the System, specifically how our application was built; it includes a demo of the website. We close the paper with a Conclusion in Section 6. In this section we mention group member contributions as well as things that could have been done differently or things that we would add if we had more time down the road. References cited throughout the paper are listed after the Conclusion in a References section.

2 Related Work

As mentioned in the Introduction, there is existing work pertaining to the upward trend in US wildfire incidence and severity over the past few decades. While there are a number of US wildfire related dashboards

* Supported by USC Viterbi School of Engineering and Professor Luciano Nocera.

in existence, we haven't found anything as comprehensive as ours that lets one interact with and walk through the data, in addition to telling a story along the way. Existing work on historical wildfire data consists mostly of exploratory and predictive analysis with simple data visualizations and maps.The other softwares we've seen simply allow the user to play around with the data on their own - however they are often not intuitive and are archaic. Our application walks the user through an interactive story, and provides them with a mission rather than simply data. The US Fire Weather Avalanche Center [7] has interactive maps which show real-time data. Unfortunately, the presentation of the data is somewhat outdated and the functionality is convoluted. While an abundance of data is provided and available, information retrieval is quite cluttered and important takeaways are not directly visible nor addressed. Our project doesn't display real-time visualizations, but provides guided visualizations that are more in-depth and clearly lay out a story through historic trends and patterns. The US Fire Weather Avalanche Center did, however, give us an initial idea surrounding how to formulate our D3 Map and what we wanted to display. In September of 2020, the LA times released a series of charts and infographics in a piece titled "The worst fire season ever. Again. [13]" This piece focused primarily on California fires, but served as a good reference and inspiration for a number of our plots. The LA times discusses the historic increase in wildfire severity and includes information on historic weather changes and wildfire suppression costs. We believe our application provides more detail and a more holistic pictures for the following reasons: we discuss the entire US; we allow the users to interact with the visualizations; in addition to suppression costs we also include prevention costs, showing the relationship between the two; we provide a deeper dive into the data through our focus on high-dimensionality. We also did research on what others have done in other areas that is still relevant. The Armed Conflict Location Event Data Project (ACLED) is a disaggregated data collection, analysis, and crisis mapping project that displays data and analysese on political violence and protests around the world [1]. We derived our thematic inspiration from their US Crisis Monitor dashboard which displays historic information surrounding various demonstration activities (COVID, BLM, etc.) across the US. Each observation is plotted as dots on a map of the US. Demonstration activities are denoted by separate colors, and demonstration sizes are encoded into the size of the dots. We determined that this was an effective way to display multivariate geographic data and used a similar approach for our D3 Fire Severity map.

3 Data

The primary data we are utilizing is from a Kaggle SQLite database referred to as the Fire Program Analysis fire-occurrence database (FPA FOD) [12]. This database contains data surrounding wildfires that occurred in the United States from 1992 to 2015. It includes 1.88 million geo-referenced wildfire records, and represents a total of 140 million acres burned during this 24 year period. The fields utilized from this database are as follows: FIRE_YEAR, FIRE_NAME, STAT_CAUSE_DESCR, LATITUDE, LONGITUDE, STATE, DISCOVERY_DATE, FIRE_SIZE. All visualizations utilize data from this source. Additionally, we utilized two supplementary datasets for a number of the visualizations. Data on US wildfire suppression and prevention funding, depicted in the two interactive wildfire spending plots comes from Congressional Research Service reports [8, 10]. Mappings for US states to their respective regions came from the US Census [4]. Please see the descriptive statistics for our dataset below:

Stat	YEAR	LAT	LON	DATE	SIZE	MONTH
Count	1.88e+06	1.88e+06	1.88e+06	1.88e+06	1.88e+06	1.88e+06
Mean	NA	3.67e+01	-9.57e+01	NA	7.45e+01	NA
Min	1992	1.79e+01	-1.78e+02	NA	1.00e-05	January
Max	2015	7.03e+01	-6.52e+01	NA	6.06e+05	December

Python [17] was utilized for all preprocessing and data cleaning performed. For the primary data source (FPA FOD) [12], we utilized sqlite3 [9] in conjunction with pandas [15] to read in the data and store it as a dataframe. Pandas was also utilized to group the data as desired for each of our six plots. When initially developing the theme for our application, we utilized matplotlib [11] to plot various cuts of the data for the purpose of uncovering insights and patterns. The python packages json and csv were utilized to export the desired cuts of data to either .json or .csv files respectively. Data from our secondary sources was minimal and as a result did not require preprocessing. This data was either manually supplemented into our datasets or copied over from extant .csv files.

4 Approach

Our approach was multifaceted in that we leveraged many concepts and tools when designing and developing our application. For the color scheme we utilized Color Brewer 2.0 [5] in order to locate a qualitative color scheme that is colorblind safe. We tested this color scheme in Chrome via the "Rendering" tool in order to ensure that the scheme accommodates all vision deficiencies. We also followed the Level AA contrast Ratio of 4.5:1 for all text throughout. We referenced Alberto Cairo's Visualisation wheel [3] in order to determine the balance of the design in each of our six plots. If judging our entire application as a singular element, it would be more in line with what Cairo describes as the wheel preferred by scientists and engineers. It favors abstraction over figuration, functionality over decoration, density over lightness, multidimensionality over unidimensionality, originality over familiarity, and novelty over redundancy. In developing each plot and the broader application as a whole, we were cognizant of visual queries. We were mindful of preattentive features such as colors, position, shape, direction, and length to present our data. As you can clearly see, our attention to these elements allows users to easily and rapidly distinguish patterns and variables. The most recent version of d3 [6] was utilized to create our plots. D3 was used in conjunction with Node.js [16], BootstrapVue [2], and Mapbox GLJS [14]. The culmination of our application was the deployment of our visuals and corresponding thematic descriptions through Vue.js. Each of our visuals was developed and tailored specifically to the goal of the information it was intended to display. The goal for each was that form follows function and that the information is conveyed in the most clear, impactful manner. The below subsections outline the specific development considerations for each of our six plots. We believe the plots and visualizations are successful in displaying their desired effect, as a result of the evaluation process we followed when designing and refining our application. In addition to leveraging development and evaluation techniques from Cairo, we sought Professor Nocera's expertise on multiple occasions. We also presented initial design ideas during one of the lecture sessions and received feedback to ensure that the plots were conveying the desired information at an ideal level of competency.

5 System

For our project, we used python to perform data pre-processing and D3 in conjunction with Node.js [16], Bootstrap Vue [2], and Mapbox GL JS [14] for our plots. Our application was deployed on Vue [18]. The application begins with a homepage that catches the reader's attention and provides them an idea of what the application is about. There are buttons at the top and bottom of the page which allow the user to navigate through the five other tabs - these tabs are as follows: Historic Fire Map, Federal Funding, Historic Fire Counts, Fire Causes. Each tab has buttons that allow one to navigate to any of the other pages within the application. In this section of the paper we discuss the critical visuals and what they represent for the purpose of the application's central goal.

Historic Fire Severity D3 Map First we want to show the fact that the severity and number of wildfires has grown in the past few decades, so we decided to plot two d3 maps side by side. As you can see in Figure 1, the left map plotted the data from the first five years in our data set, and the right one the last five years. This was done to allow users to better understand the juxtaposition in fire incidence and severity between the beginning and ending of the period. With symbol size proportional to the size of wildfires, one would quickly notice that the wildfires have been more prevalent and severe in a split second. The users are able to click on a specific fire cause in order to explore how that sole cause changed over the period. This visual shows that there has been a large increase in both fire instance an severity from the beginning of the 23 year period to the end.

Funding Plots We then wanted to inform the users of the historic trend in wildfire funding and how it can be broken down. As you can see in Figure 2, we chose to overlay a breakdown of suppression vs prevention

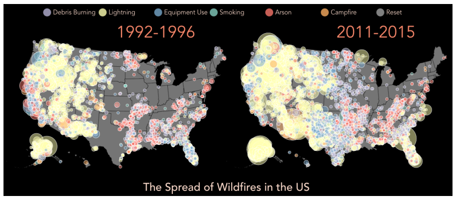

Fig. 1

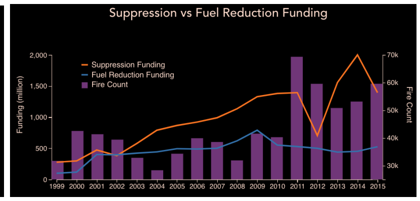

Fig. 2

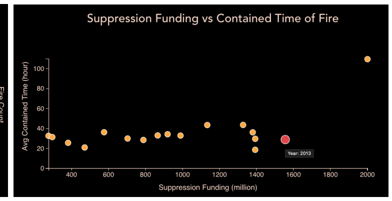

Fig. 3

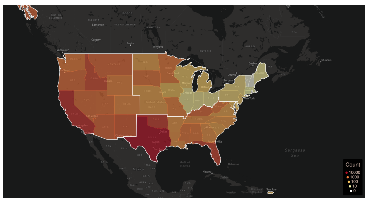

Fig. 4

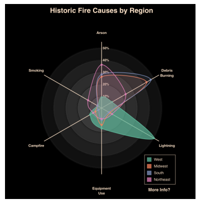

Fig. 5

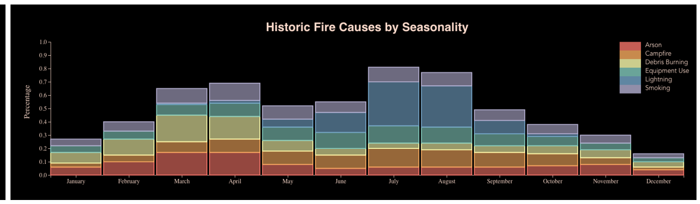

Fig. 6

funding over a bar chart of fire counts. It is clear that while suppression funding has increased over the period, funding for fire prevention has remained stagnant. Fire suppression funding has increased heavily with an increase in fire instance. We develop the concept further in Figure 3 by showing that there is little-to-no correlation or relationship between suppression funding and containment time. This suggests that increase in suppression funding does not have a direct impact on the containment time of fire, and drives home our point that additional funding for fire suppression might be better allocated to prevention instead.

Fire Counts Mapbox Map After showing how suppression might not be the most effective way of dealing with wildfires, we then start to explore different dimensions of the wildfires as an attempt to come up with practical suggestions. First, the regional factor of the wildfires is explored by a choropleth map in mapbox map (Figure 4) and a radar chart (Figure 5). Mapbox allows users to zoom in and out if they want to focus on specific areas (maybe their hometown) to see how much that area has historically suffered.

Fire Causes Plots The radar chart serves as a clear summary of the regionality of the various wildfire causes. Depicts the proportion each fire cause contributed to the total fire count, by region. For example: over the period, 50 percent of the fires in the West were caused by lightning, and 17 percent were caused by Equipment use. This plot shows the most common wildfire causes in the various regions of the US and shows that while there is overlap between certain regions, each region has specific causes that are more common: the West has strictly lightning to worry about; the Northeast could cut down on its wildfires significantly if it focused first on Arson and then on Debris burning; the South and Midwest should focus on Debris Burning and Arson. The radar chart also allows allows for interaction. If you hover over one of the stars/blobs, the area for that region of the US will become highlighted. Also, hovering the pointer over a data point will reveal the value of that point. We finish by exploring the seasonality of wildfire causes with a stacked bar chart. As you can see in Figure 6, the stacked bar chart shows 6 most common wildfire causes in the US and their prevelance during different months. In general, arson and debris burning take up the largest portion of wildfires in Spring (March and April) while lightning strikes and campfire are two main threats during summer days (July and August). But lightning rarely causes wildfires in other seasons. Smoking and equipment use, on the other hand, show a relatively consistent rate in the whole year. When the browser window is small enough, the stacked bar chart would be simplified as a 'stacked' lollipop chart where the original rectangles are simplified as lines and circles.

Final Tab The final tab is a conclusion slide that summarizes and brings together our theme and the story our application walks the user through. This tab also includes a link to our project repository for any interested.

6 Conclusion

Fires in the US have been increasing in size and instance for several decades. As seen through our application, we need to work toward preventing wildfires as opposed to simply suppressing them. The application we created will be useful for both the private and public sectors. Government agencies and fire authorities, as well as the general public, will be able to utilize the information presented to better identify situational risks and target activities that are more likely to result in sparking a wildfire. We have contributed to society an actionable mission via acutely summarized data and a guided progression through interactive visuals. We broke up the tasks by what drew the various team members' interests. Po set up the Sozi presentation for the proposal and final presentation, performed initial exploratory data analysis for determining the goal of the project with Luke, managed the initial setup of Vue with Asumi, and created the Historic D3 Fire Map. Asumi set up the group Slack channel, managed the initial setup of the Vue app with Po, developed the two Federal Funding charts, and conducted the Bootstrap implementation. Luke performed the initial exploratory data processing and analysis for determining the goal of the project with Po, crafted the Historic Fire Causes radar chart, conducted the data pre-processing for the responsive plot, and wrote the essay/paper. Kaile located the primary wildfires dataset, and created the Fire Causes responsive bar chart as well as the Mapbox map. All team members contributed to data pre-processing and grouping for their respective plots, talking points and picture selection for the sozi presentations, and creation/presentation of the video. If we had more time or were to continue this as a future work, we would first like to locate additional data to further prove the notion that more attention should be given to preventing wildfires as opposed to suppressing them. This could include data from studies that show fire prevention is effective and viable from a funding perspective. We would also like to include information that breaks down costs and processes for prevention of various fire causes. A design element from Cairo's Visualization Wheel we would like to add if we had more time would be to include more decoration within our application. This could include elements in the margins that create a boom effect to attract the reader and to help with memory of certain concepts or themes. We would additionally like to show data in a higher dimensionality. For instance, we displayed a plot that shows the proportion of fire causes by region, and one that depicts the breakdown of fire causes by month of the year, but would like to create another plot that incorporates all three dimensions: fire cause, month of the year, region. We think it is possible that observing the data in such a manner could yield further important insights. This was an incredibly rewarding project and as a team, we are very please with the results and the amount that we learned throughout the process.

References

1. Us crisis monitor. https://acled.github.io/US/, last accessed 15 Nov 2020
2. Bootstrap. https://getbootstrap.com/, last accessed 21 Nov 2020
3. Cairo, A.: The Functional Art: An introduction to information graphics and visualization. New Riders (2012)
4. Census regions and divisions of the united states. https://www.census.gov/geo/pdfs/maps-data/maps/, last accessed 21 Nov 2020
5. Colorbrewer 2.0: Color advice for cartography. https://colorbrewer2.org/type=sequentialscheme=BuGnn=3, last accessed 21 Nov 2020
6. d3js. http://d3js.org, last accessed 21 Nov 2020
7. U.s. wildfire map — fire, weather avalanche center. https://www.fireweatheravalanche.org/fire/, last accessed 18 Nov 2020
8. Gorte, R.W.: Federal Funding for Wildfire Control and Management. Congressional Research Service (2011)
9. Hipp, R.D.: SQLite (2020), https://www.sqlite.org/index.html
10. Hoover, K.: Federal Wildfire Management: Ten-Year Funding Trends and Issues (FY2011-FY2020). Congressional Research Service (2020)
11. Hunter, J.D.: Matplotlib: A 2d graphics environment. Computing in science & engineering **9**(3), 90–95 (2007)
12. 1.88 million us wildfires. https://www.kaggle.com/rtatman/188-million-us-wildfires, last accessed 21 Nov 2020
13. Krishnakumar, P., Kannan, S.: The worst fire season ever. again. Los Angeles Times (2020)
14. mapbox. https://www.mapbox.com/, last accessed 21 Nov 2020
15. McKinney, W., et al.: Data structures for statistical computing in python. In: Proceedings of the 9th Python in Science Conference. vol. 445, pp. 51–56. Austin, TX (2010)
16. Openjs foundation: Node.js. https://nodejs.org/en/, last accessed 21 Nov 2020
17. Van Rossum, G., Drake, F.L.: Python 3 Reference Manual. CreateSpace, Scotts Valley, CA (2009)
18. Vue.js. https://vuejs.org/v2/guide/, last accessed 21 Nov 2020