How have wildfires in California changed over time? The potential implications for future wildfire management and preparedness.

Lee Hong Teng
01385472
htlee.2021@scis.smu.edu.sg

Foo Chuan Wei
01394670
cwfoo.2021@economics.smu.edu.sg

Tan Jing Xiang
01422949
jxtan.2021@scis.smu.edu.sg

Edlyn Tan Tse Yin
01444466
edlyn.tan.2021@scis.smu.edu.sg

Angela Camp Torres
01522978
angela.ct.2024@exchange.smu.edu.sg

1 INTRODUCTION

A wildfire is an uncontrolled fire that can spread through forests, grasslands, savannas, and other ecosystems. Additionally, they are not constrained to any area or environment. (National Geographic Society, 2025). An example of a wildfire would be the 2018 CampFire with Butte County California, where most of the town was destroyed, this resulted in 86 deaths and approximately 18,804 structures burnt (*Camp Fire*, n.d.).

Wildfires will threaten lives, destroy homes, and leave lasting scars on the environment and are seen more prominently in regions like California. California's wildfires have become more frequent and destructive each year, with fifteen of California's twenty largest wildfires happening since 2000 (Science: Wildfire Impacts, n.d.).

Climate change is making things worse by creating drier, hotter conditions that make fires easier to start and harder to contain (MacCarthy, n.d.). The 2025 fire season is already showing worrying signs—by February 3, there had been 337 wildfires, scorching 57,636 acres. That's a massive leap from the same time in 2024, when just 141 fires burned a mere 76 acres (CAL FIRE, n.d.). Our project involves making several visualizations with past datasets of California wildfires, so that we can identify patterns and trends that can help lower risk or damages of wildfires.

2 MOTIVATION

California's wildfire crisis is driven by multiple factors such as climate change, land management practices, and fire suppression policies. With weather becoming hotter and dryer it makes it easier for fires to ignite and spread. At the same time, decades of fire suppression and the decline of Indigenous fire practices have led to an accumulation of dry fuel, making wildfires even more destructive (Drivers of California's Changing Wildfires: A State-of-the-Knowledge Synthesis, n.d.). The mentioned factors are non-exhaustive. Therefore, to better understand and communicate trends, open-source data plays an important role.

With so many cases of damages, loss of lives and homes, by referring to the case studies in California, we might be able to learn from them to avert or reduce the impacts of wildfires in the future.

By analyzing wildfire data, our group aims to uncover patterns, track changes over time, and visualize the impact of different contributing factors. Such insight can be applied globally as well, and can help policymakers and fire management agencies develop more effective strategies to prevent and mitigate wildfires, ultimately reducing their impact on communities and the environment (MacCarthy, n.d.).

3 RELATED WORKS

3.1 Existing Dashboards

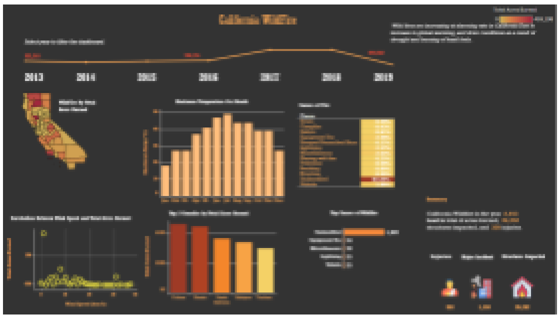


Figure 1: Interactive Dashboard on California Wildfires by Taravat (Tableau)

Through interaction with this Tableau dashboard on California Wildfires (Taravat, 2022) in Figure 1, we are able to filter data by years through 2013-2019 to gain valuable insights into California wildfires. The dashboard includes a choropleth map displaying total acres burned per county, complemented by a bar chart highlighting the top counties most affected by fire. Additional visualizations include a bar chart showing maximum temperatures per month, a scatterplot illustrating the correlation between wind speed and total acres burned, and a breakdown of various fire causes, supported by a bar chart of the top five causes by count. There is also a count of the total number of injuries and destroyed structures. Through this dashboard, we are able to identify spatial and temporal trends, understand key contributing factors to wildfire severity, and make data-driven observations that could support prevention and mitigation strategies.

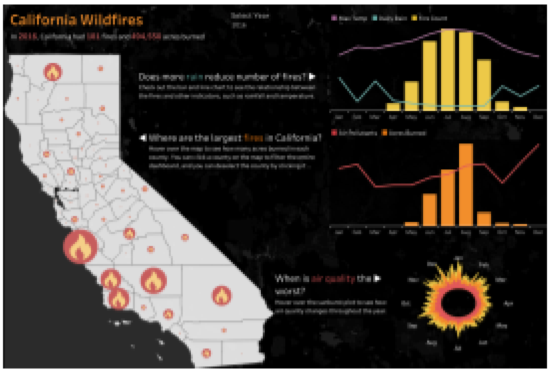


Figure 2: Interactive Dashboard on California Wildfires by Sandeep Tiwari (Tableau)

Similarly, in another dashboard on California Wildfires (Tiwari, 2022) as shown in Figure 2, we are able to filter by county and years through 2003-2019 to explore data related to the California wildfires. The dashboard encompasses visualisations such as a choropleth map indicating wildfire intensity across counties (by acres). Other visualisations include mapping correlations between maximum temperature and daily rain to fire count, and the concentration of air pollutants to the number of acres burned. A sunburst plot is also included in the dashboard to visualise how air quality varies throughout the year. By engaging with this dashboard, whether variables like temperature, rainfall and concentration of air pollutants impact the quantity and severity of wildfires can be analysed.

3.2 Key Issues

(1) Outdated Data

While these dashboards provide valuable insights into the correlations between potential causes of wildfires, the data utilized only extends up to 2019. In light of the increasingly rapid changes in the Earth’s climate in recent years due to factors such as Global Warming, the usage of more recent data in visualisations is essential for accurately understanding the evolving patterns and underlying causes of wildfire occurrences.

(2) Scale of Visualisations

The first dashboard mainly focused on the general causes of wildfires while the goal of the second dashboard was to find a correlation between selected variables and the occurrence of wildfires. We feel that there is a need to explore both the general causes as well as specific variables. This is because there are a large number of wildfires with ‘Unidentified’ causes – which could likely be mapped to variables such as Temperature and Rainfall.

4 DATASETS

We will use multiple wildfire datasets to examine trends in wildfire incidents, severity and causes within the California state and its impacts on cities within.

4.1. Wildfires Data Sources

- **California Fire Perimeters**

  Annually updated data that provides geospatial data on wildfires perimeters, severity, and causes up till 2023 on both public and private lands

- **CA Perimeters NIFC FIRIS**

  Live Data of Fires from 2024 till current. It only has data on where the fire is and its severity but no knowledge on the direct cause. We use this information still as it can give a complete picture to our dataset

4.2. Climate and Geographical Sources

In addition to wildfire incident data, climate will be incorporated to identify and visualise relationships between nature and the occurrences of wildfire

- **TIGER/Line Shapefiles**

  The US county files were used to allow us to analyse wildfires on a county level within the state of california

- **NOAA Climate Data**

  This dataset provides monthly Temperature data for california across the United States, covering the period from 1895 to 2025

- **California Weather and Fire Prediction**

  This dataset provides Weather observations and wildfire data in California. Specifically we will be using the precipitation data (Total monthly precipitation levels, measured in inches) and wind speed

5 TOOLS AND RESOURCES

5.1 Tableau

Tableau was used for our initial exploration of datasets, in addition to creating prototypes for our visualizations. We were able to merge our datasets and form many different visualisations during our exploration. By exporting the files as a workbook, we were able to share it with one another to assist with collaboration and to add on to our research. In the end, we used tableau for map visualisation to utilise our geographic data.

5.2 D3.js + React

React was used for the development of the frontend for the web application. It lets us have a responsive and modular user interface, allowing for convenient viewing by our user.

D3.js is a Javascript library that allows us to make a wide variety of visualizations. Our final visualizations were done on

D3.js, as it allowed for us to build a more customized dashboard as compared to Tableau, allowing for better interaction for the users.

5.3 Node.js and Express (Backend)

Node.js allows for us to create servers, web apps, command line tools and scripts as needed, allowing us to host our visualizations on the web application when necessary.

Express is a minimal Node.js framework that we have decided to use due to it being lightweight and having good performance for our project.

5.4 CSV-Parser, Shapefile and Turf.js (Data Processing)

CSV-parser was used to convert our csv files to JSON files quickly, so that we can consolidate and use the datasets we have downloaded for our D3 visualizations.

Shapefiles are used for storing geometric locations and attribute information of geographic features. The shapes files are used to more easily process and read our data.

Turf.js was used as a geospatial analysis library. It was important in helping us handle spatial computations according to the data.

5.5 Docker (Containerization)

Docker and docker compose were implemented mainly to simplify configuration. As mentioned we used D3.js and React for frontend and Node.js for backend configuration.

5.6 Python (Pandas)

Pandas was used to process the US county shape file by filtering out the counties in California instead. The data was also duplicated by periods of years for joining reasons in tableau.

6 TASKS

To achieve our goal of identifying patterns and trends that could potentially lower the risk and damages of wildfires, we have identified 5 research areas as shown in Table 1 below, with tasks assigned for each one.

Research Area	**Purpose**
Visualise seasonality trends of wildfires	Identify specific time periods to pay extra attention to containing wildfire in the region
Understand general trends of causes of fire	Identify any important factors that could have led to the majority of these wildfires

Understand potential correlations between climate variables and wildfire count/severity	Inspect whether there are any direct causal relationships to wildfires
Analyse causes of fire in counties in california	Ensure the provision of more targeted fire prevention measures

Table 1: Research Areas and corresponding Purposes

6.1 Identify seasonal trends of wildfires in recent years

Recognising periods of peak wildfire spreading allows for advanced risk management and more informed policy planning.

This comes in the form of increased resources being allocated to the prevention of wildfires, and advanced land use policies and insurance practices.

6.2 Identify general trends of wildfire causes

By pinpointing the common causes of wildfire, targeted measures can be taken to address them, preventing more wildfires that could be potentially caused. This allows for more efficient resource allocation as well.

For example, laws can be tightened for outdoor burning and lightning protection measures can be prioritised.

6.3 Identify correlations between climate variables and the occurrence of wildfires

Visualisations on how the climate is linked to the occurrence of wildfires allows for a more reliable forecast of fire danger.

This is useful as advanced warnings would result in better preparedness during high-risk periods.

6.5 Analyse the worst fire in recent years

Zooming into the worst fire that occurred in recent years. This provides critical insights into which factors contributed to its severity, allows us to discover limitations and recognise how similar disasters can be prevented in the future.

Lessons learned from this analysis can help to identify improvements in emergency planning and resource allocation.

7 METHOD

Tableau

Tableau was used by our group to sketch out initial ideas of visualizations to be put in D3. We faced issues with loading our visualizations due to our data size and curbed this by creating extracts before running our visualizations.

D3.js

As the data files came in CSV, Shapely and Geojson formats. We converted them all into Geojson for standardisation. Afterwhich, we found out that Geojson files were too large due to the polygon data that we were not using, causing it to take too much time to process the files to display them. As such we extracted the data we needed visualisation into json format.

8 VISUAL ANALYSIS RESULTS & DISCUSSIONS

Our visual analysis of California wildfire data has led to several significant findings across different dimensions of the dataset. This section outlines the major observations derived from the analysis, organized by the task, and discusses their implications. Each insight is supported by relevant visualizations with annotated figures to validate our conclusions. A bigger picture of each figure may be found in the appendix

8.1 Seasonality of wildfires

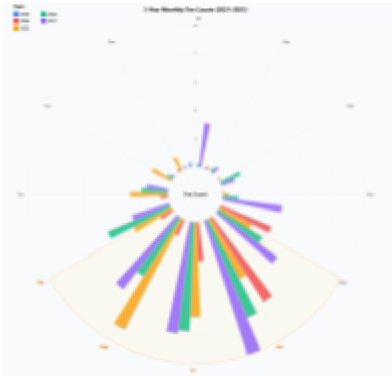


Figure 3: Radial Chart of Fire Counts (2021-2025)

One of the most striking findings from our analysis concerns the seasonality of wildfires in California. As seen in the radial bar chart (Figure 3), which visualizes monthly fire counts over the past five years (2021-2025), there is a clear and consistent peak during the summer months, particularly from June to September. Among these, August stands out with the highest number of fire incidents in 2023. The visual impact of the longer bars during this period highlights a strong seasonal trend that aligns with California's hottest and driest weather. This reinforces the need for heightened prevention efforts and emergency preparedness during these months, specially in August, when wildfire risk is at its peak. Such seasonality insights are crucial for informing policy and optimizing the timing of public awareness campaigns, firefight resource allocation and landscape management initiatives.

8.2 General Trend of Fire Causes

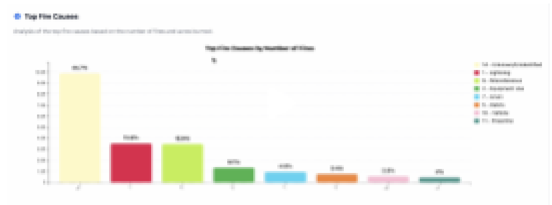


Figure 4: Top Fire Causes

In terms of fire causes in 2023, our analysis reveals that the majority of fires are attributed either to “Unknown” causes or Lightning, as shown in the bar chart of fire causes *(Figure 4)*. This is largely consistent across multiple years, and although lightning is a naturally occurring factor, the large proportion of “Unknown” causes is problematic. “Unknown” may arise when fire investigators are not sure about the original cause behind the fire (U.S Department of the Interior, n.d.)It reflects a limitation in the reporting process, which ultimately hinders the development of more targeted prevention policies. Further efforts in fire investigation and classification could significantly improve future wildfire response and management.

8.3 Importance of Climate Variables

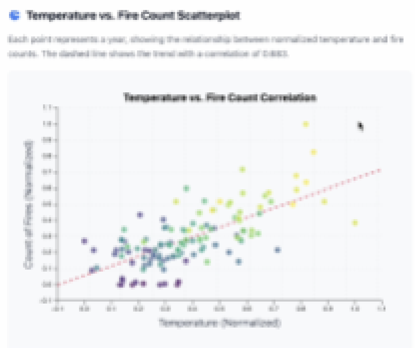


Figure 5: Temperature vs Fire Count Scatterplot

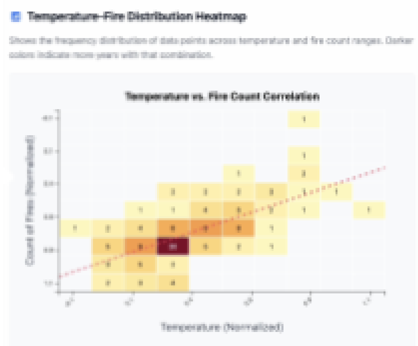


Figure 6: Temperature-Fire Distribution Heatmap

The role of climate variables, particularly temperature, is further confirmed through a correlation analysis. As demonstrated in the scatterplot and heatmap (Figure 4 and 6), there is a strong positive correlation (r = 0.683) between temperature and the number of fires. This statistically significant relationship supports the hypothesis that hotter conditions directly contribute to increased fire activity. This insight not only confirms existing climate science but also strengthens the case for incorporating weather and temperature trends into fire risk models.

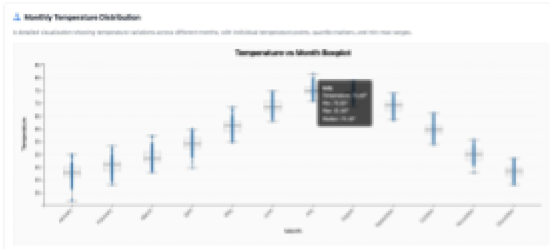


Figure 7: Monthly Temperature Distribution

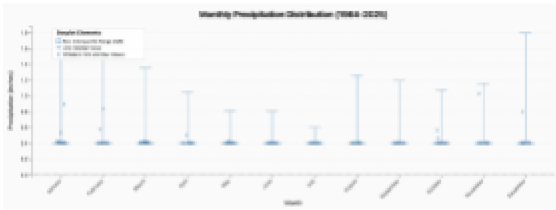


Figure 8: Monthly Precipitation Distribution

The weather conditions in California are as such: We observe temperature rising during the summer as expected with July being California's hottest month as shown in Figure 7. The precipitation data in Figure 8 has been scaled via logarithm as precipitation in california is very small across the board.

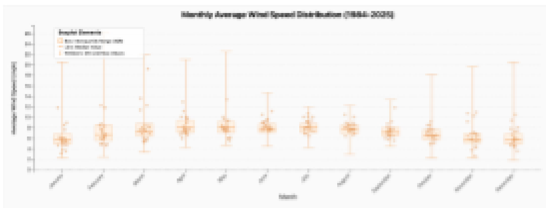


Figure 9: Monthly Average Wind Speed Distribution

Precipitation in California is minimal but we conclude based on the distribution that rainfall is especially the lowest in the summer since the higher tail curves down in the summer before going up again. In terms of the monthly wind speed distribution based on Figure 9 it may look like the wind speed picks up in the summer but we noted that this is so as the interquartile range has shrunk. In terms of a comparison of medians, our group finds that we could not really establish a clear relationship between wind speed and fire count as the median wind speed are pretty constant across the years. Theoretically these three weather trends create an optimal condition for wildfires to occur especially during summer as a combination of lesser precipitation and hotter temperatures results in drier lands which will make vegetation easier to catch fire. Higher wind speeds would mean that it is easier to catch fire and also spread fire quickly. However our windspeed box plot distribution is contrary to our hypothesis, Wind speeds did not have a strong & clear relationship to increasing fire count . Nevertheless the other two climate variables are inline with the seasonality of wildfires in California. Hence summer is the wildfire season in California where they experience more wildfires.

8.4 Analysing the worst fire from 2022 2024 & 2025

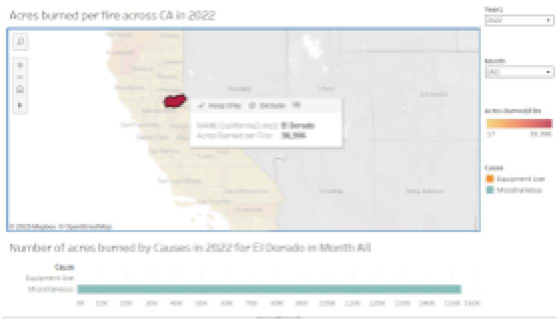


Figure 10: Worst Fire in 2022 - El Dorado

In conjunction with our goals, our group attempts to explain some of the causes of the recent worst wildfires in California history. We use Acres burned per fire as it gives a more proportional understanding on the severity of the fire. As per figure 10, 2022 saw the worst fires in El Dorado. The direct cause of fire attributed to El Dorado happened to be Miscellaneous of which it takes up a large proportion of Unlike the ‘Unknown’ , Miscellaneous fires are causes of fire that encompasses a wide range of causes that does not fall under a single category (U.S Department of the Interior, n.d.). For example these fires may arise from spontaneous combustion or specifically powerlines trips in this case (YubaNet, 2022).

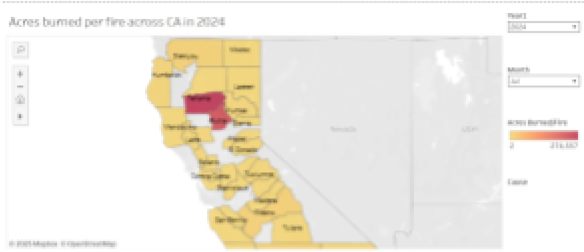


Figure 11: Worst Fire in 2024 - Tehama & Butte

Our causes of fire dataset only gives us data till 2023, however this does not deter our group from explaining potential causes of fire in 2024 onwards. As per figure 11, the worst fires in 2024 happened in Tehama and its neighbour Butte. These fires happen in July as well which coincides in the summer period where wildfires are more prevalent with the prime conditions mentioned previously.

9 CONCLUSION & FUTURE WORK

Through our analysis of California wildfires and their potential causes, we have identified some limitations that we faced, and have several recommendations on how our analysis can be used for future improvements.

9.1 Limitations

Even though we made the most of the datasets that were available, we were faced with the following limitations:

(1) Ambiguous classification of wildfire causes

A large number of fires were categorised as "Unknown" or "Miscellaneous," limiting the ability to accurately determine and address the primary causes of wildfire. With more accurate classification, more insights could have been derived from our analysis.

(2) Single-State Focus (California)

While we chose to focus our analysis on California, a hotspot for wildfires, our findings may not apply to other regions or countries with different ecological or climatic conditions. For example, an increasing number of wildfires have been reported in South Korea. However, due to climate differences, the insights that we have obtained from the California wildfires may not be applicable to them.

9.2 Future Work

Through our analysis, we were able to gain crucial insights which have led us to make the following recommendations.

(1) Efficient Resource Allocation

Through our analysis, common causes of the California wildfires and their seasonal patterns have been identified. Through this, we suggest a more efficient and strategic use of resources, in the form of manpower, equipment and budget. Targeted prevention campaigns and infrastructure improvements can also be introduced, which will reduce response time and minimise damage. For example, treatment burns are used to prevent large catastrophes from happening. It may be prudent to execute treatment burns before the wildfire season in summer that we have pinpointed out

(2) Enhance Fire Cause Classification Efforts

To address the limitation of the ambiguous classification of the causes of wildfires, we suggest a partnership with relevant agencies to improve cause reporting. Through better classification of the causes of wildfires instead of grouping them as 'Miscellaneous', prevention strategies can be improved.

(3) Develop Predictive Models

Through the insights derived from our analysis, Machine Learning models can be further applied to build predictive models for wildfire occurrence. This helps to better predict potential wildfires that may occur, leading to faster prevention and response efforts.

Appendix

Figure 3: Radial Chart of Fire Counts

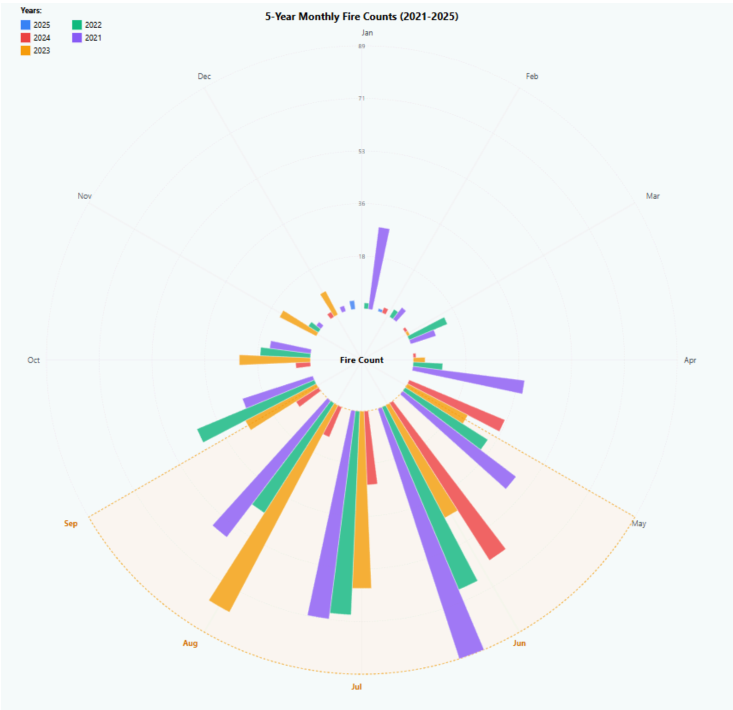


Figure 4: Top Fire causes

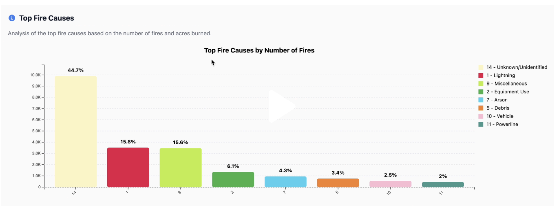


Figure 5: Temperature vs Fire Count Scatterplot

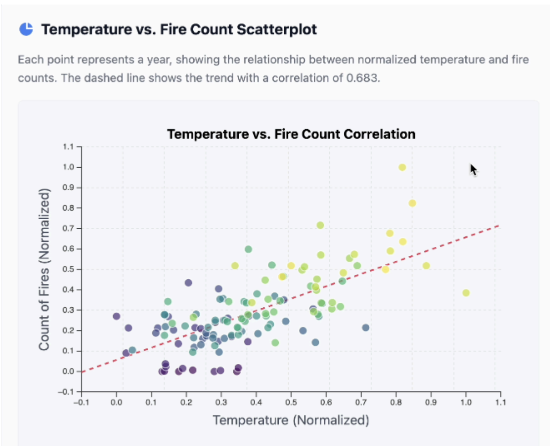


Figure 6: Temperature-Fire Distribution Heatmap

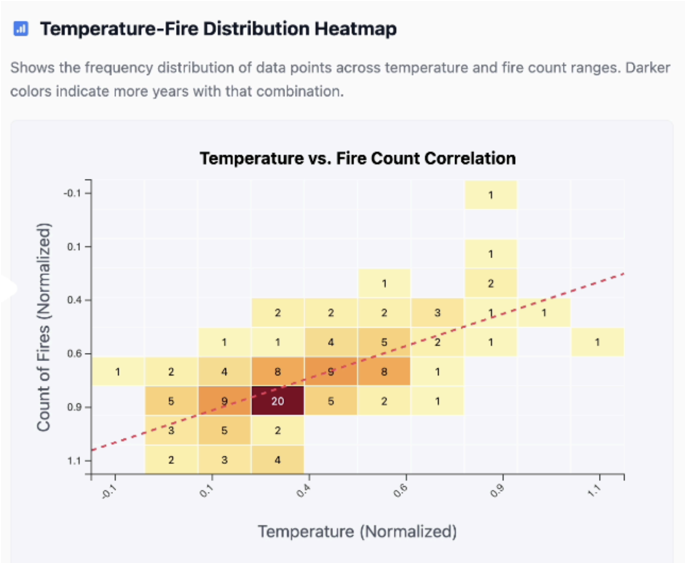


Figure 7: Monthly Temperature boxplot

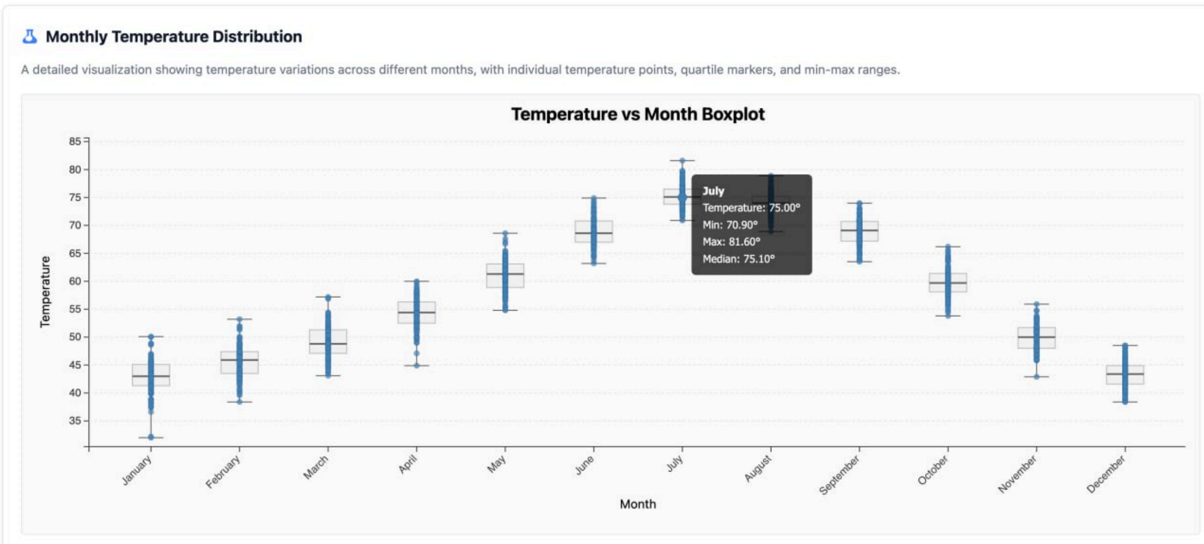


Figure 8: Monthly Precipitation boxplot

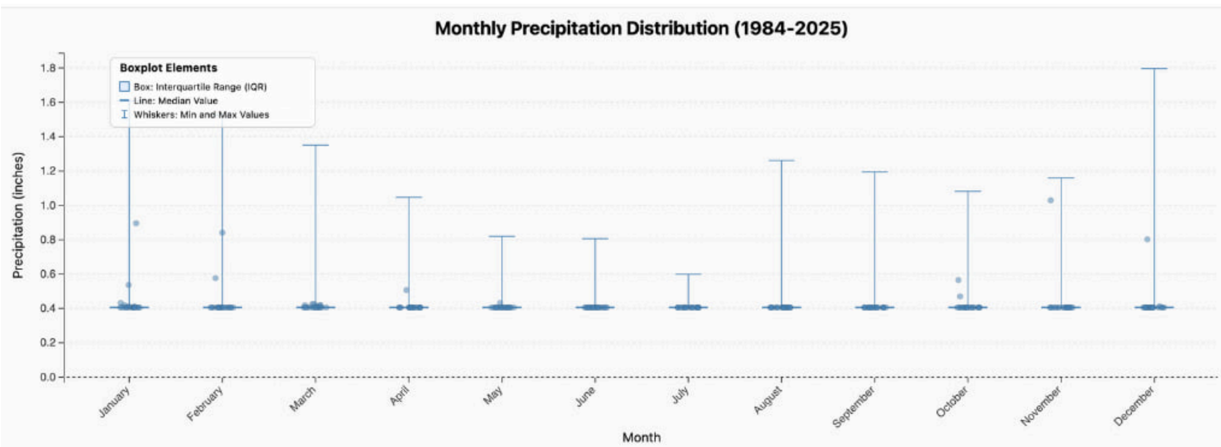


Figure 9: Monthly Wind Speed boxplot

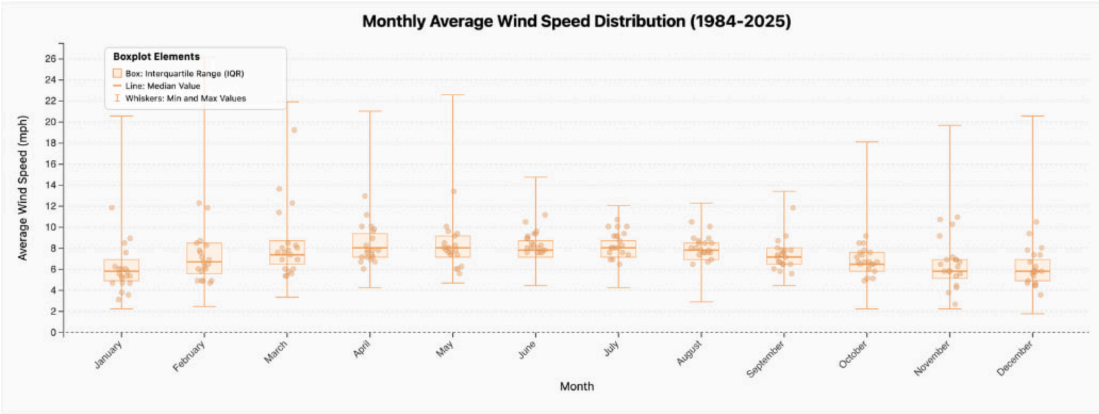


Figure 10: Worst Fire in 2022

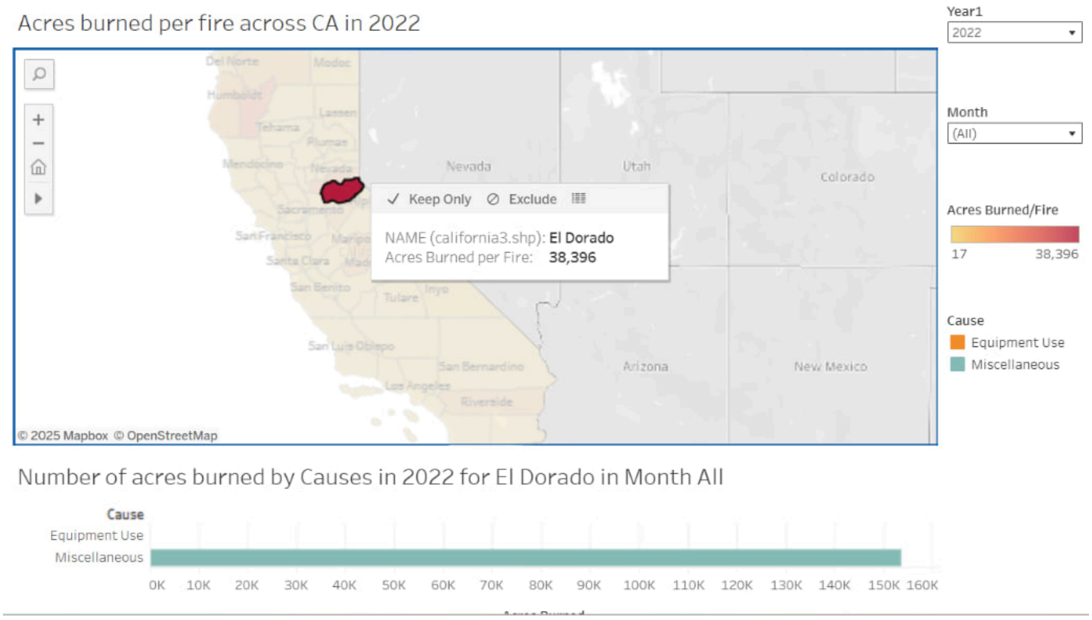
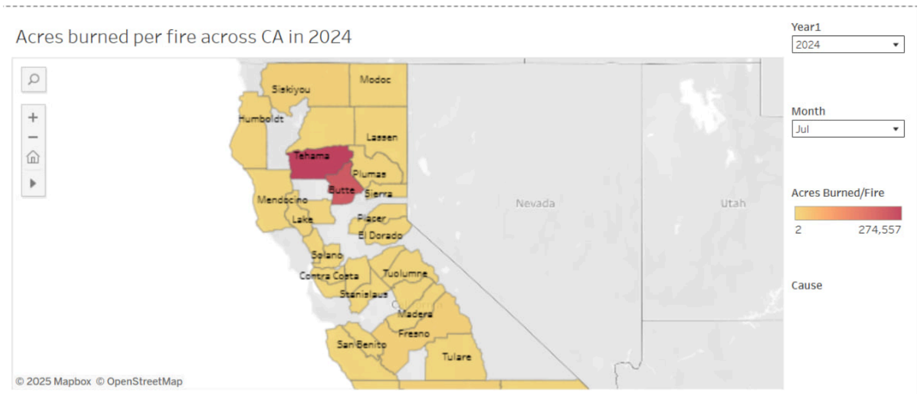


Figure 11: Worst Fire in 2024



References

Camp Fire. (n.d.). Butte County. Retrieved April 5, 2025, from https://www.buttecounty.net/342/Camp-Fire

Fire Perimeters | CAL FIRE. (n.d.). Cal Fire. Retrieved April 5, 2025, from https://www.fire.ca.gov/what-we-do/fire-resource-assessment-program/fire-perimeters

Land Cover Data Overview | U.S. Geological Survey. (2019, February 13). USGS.gov. Retrieved April 5, 2025, from https://www.usgs.gov/programs/gap-analysis-project/science/land-cover-data-overview

MacCarthy, J., Richter, J., & Karzen, K. (2025, February 5). *4 Graphics Explain LA's Rare January Fires*. World Resources Institute. Retrieved April 5, 2025, from https://www.wri.org/insights/los-angeles-fires-january-2025-explained

National Centers for Environmental Information | National Oceanic and Atmospheric Administration. (2025, March). *NOAA National Centers for Environmental information, Climate at a Glance: Statewide Time Series*. Retrieved April, 2025, from https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/statewide/time-series/1/pcp/12/1/2020-2025?base_prd=true&begbaseyear=2020&endbaseyear=2025&trend=true&trend_base=10&begtrendyear=1895&endtrendyear=2025

National Geographic Society. (2025, January 15). *Wildfires*. National Geographic Education. Retrieved April 5, 2025, from https://education.nationalgeographic.org/resource/wildfires/

RDI.CAL FIRE. (2024). *CA perimeters NIFC Firis public view - dataset - California open data.*

Science: Wildfire Impacts. (n.d.). California Department of Fish and Wildlife. Retrieved April 5, 2025, from https://wildlife.ca.gov/Science-Institute/Wildfire-Impacts

Southwest Climate Adaptation Science Center | SW CASC. (n.d.). *Drivers of California's Changing Wildfires: A State-of-the-Knowledge Synthesis*. Drivers of California's Changing Wildfires: A State-of-the-Knowledge Synthesis

Statistics | CAL FIRE. (n.d.). Cal Fire. Retrieved April 5, 2025, from https://www.fire.ca.gov/our-impact/statistics

Taravat. (2022, December 9). *California Wildfires*. Tableau. https://public.tableau.com/app/profile/taravat/viz/CaliforniaWildfires1/Dashboard1

The California Wildfire Data. (n.d.). Kaggle. Retrieved April 5, 2025, from https://www.kaggle.com/datasets/vijayveersingh/the-california-wildfire-data/data

Tiwari, S. (2022, December 3). *California Wildfires*. Tableau. https://public.tableau.com/app/profile/sandeeptiwari/viz/CaliforniaWildfires_16074146008400/CaliforniaWildfires

U.S. Department of the Interior (Trans.). (n.d.). *Wildfire Investigations | Indian Affairs*. BIA. Retrieved April 7, 2025, from https://www.bia.gov/service/wildfire-prevention/wildfire-investigations

YubaNet. (2022, September 8). *PG&E submits Electrical Incident Report re The mosquito fire*.https://yubanet.com/regional/pge-submits-electrical-incident-report-re-the-mosquito-fire/