Analyzing California Wildfires (2005-2015): A Comprehensive Study Using Spatial Data and Visualization

Renqi Huang[1], Jade Hacket[1], Jiaqi Liu[1], and Tianliang Tao[1]

University of Southern California, Los Angeles CA 90007, USA
{renqihua, jliu9100, sjhacket, taotianl}@usc.edu

Abstract. Wildfires occur frequently in California, and they pose a significant threat to human life and the environment.In this project, we used two datasets to analyze wildfires in California from 2005 to 2015. Our goal was to create an informative dashboard to visualize key aspects of these wildfires. The dashboard focuses on causes, cumulative fire size by county, the number of fires by county, the largest fires by decade, and a predictive map highlighting areas at risk of future wildfires. This research will enhance our understanding of past wildfires and provide strategies for containing and preventing wildfires in the future.

Keywords: Wildfires · D3 · Visualization · Pre-processing.

1 Introduction

In recent years, wildfires in California have increased in frequency and destruction and have become a major problem. We will delve into California wildfire data with the goal of revealing the causes and spatial distribution of wildfires and providing valuable information and strategies for preventing and containing wildfires in the future. The rising global relevance of wildfires underscores the urgency of this study. The motivation for this research is rooted in the urgent need to comprehensively address the consequences of California wildfires. These events pose a direct threat to human lives and ecosystems, and their effects extend far beyond immediate damage, impacting the state's environmental, economic, and public health realms. Effective communication and understanding of wildfire dynamics are essential in this context. Our study utilizes visualization to transform complex wildfire data into intuitive, actionable insights, catering to a wide audience from policymakers to the general public.

This paper's scope spans a wide-ranging reader, emphasizing the necessity of making wildfire information accessible and comprehensible to diverse stakeholders. Our approach addresses the complexity of wildfire data, incorporating multifaceted factors like climate, vegetation, and human activities, and presents this information through an advanced visualization dashboard. This dashboard not only depicts the current state of wildfires but also explores their evolving

nature over time. We will use Alberto Cairo's visualization wheel to improve the visualization for California wildfire analyzing.

Our work stands out for its comprehensive and integrated approach to studying California wildfires. Unlike previous research focusing on isolated aspects, our study combines various data dimensions – causes, fire size, frequency, and predictive mapping – into a singular, user-friendly visualization tool. This holistic view offers new insights into the interplay of different factors affecting wildfires.

The paper is structured to guide readers through our research journey. In section 2,we review related wildfire visualization works. In section 3, we describe our datasets and its preprocessing. In section 4 We discuss the design considerations of our visualization tool. In section 5, we demonstrate how to built the dashboard. In section 6, We conclude by summarizing our contributions and proposing future research directions.

Essentially, this study is an important exploration of wildfires in California, providing practical visualization and analysis. Our goal is to deepen understanding and develop effective strategies to mitigate the impacts of wildfires, leading to a better future for California communities and ecosystems.

2 Related Work

Multiple studies in the field of wildfire research, particularly in visualization and predictive analytics, have laid the groundwork for our current understanding and methods. A key focus area has been the causes and spatial distribution of wildfires.The Statista created a bar chart shows that the average number of acres burned by wildfires in California is growing each year. [2] The visualization of the number of fires by county has been another critical area. Research efforts by Dong et al. incorporated geographical information systems to map fire regimes. [3] This study provided foundational insights into spatial distribution but often did not offer the dynamic, interactive visualization tools.

Our work aims to fill these gaps by not only analyzing historical data, but also incorporating predictive modeling to identify areas at risk for future wildfires. This project offers great potential for prevention and mitigation strategies. By integrating data on wildfire causes, cumulative fire size, frequency, and predictive mapping into a comprehensive visualization dashboard, our study provides a holistic view of wildfires in California. This multifaceted approach allows for a more nuanced understanding of the interactions between the different factors affecting wildfires, a critical step forward for the field of wildfire management and research.

3 Data

Our analysis is anchored on two datasets: the Burned Area Boundaries Dataset from MTBS, and the 1.88 Million US Wildfires from Kaggle. The MTBS dataset, a collaborative effort between the U.S. Geological Survey and the USDA Forest Service, maps the perimeters of significant wildland fires across the U.S. from

1984 to 2022. [1] It provides detailed data on fire locations, sizes, and severity. The Kaggle dataset, spanning 1992 to 2015, comprises a comprehensive record of U.S. wildfire incidents, including specifics like fire cause, discovery date, and containment details. [4]

In terms of pre-processing, we used Numpy and the Pandas package in Jupyter Notebooks to address missing values and inconsistencies in both datasets. For the MTBS data, this meant refining the fire perimeter records to ensure accuracy in boundary delineation. In the Kaggle dataset, we focused on standardizing the entries for fire causes and dates. We also separate these two datasets into mutiple json file, each json file match the one chart in the dashboard. We aligned them based on geographical and temporal parameters, ensuring that the fire incidents in the Kaggle dataset corresponded accurately with the fire perimeters in the MTBS data.

4 Design

Combined with the concept of the visualization wheel, our design process carefully balances complexity and readability. We prioritized multidimensionality and functionality to ensure that our visualizations effectively communicate the complexity of wildfire data. Design considerations were made to balance novelty, user appeal, and familiarity to ensure intuitive understanding. Our visualizations stand out because they emphasize graphic clarity over decoration, ensuring that each design element serves a purpose. Figure 1, shown below, shows the visualization wheel for the dashboard design.

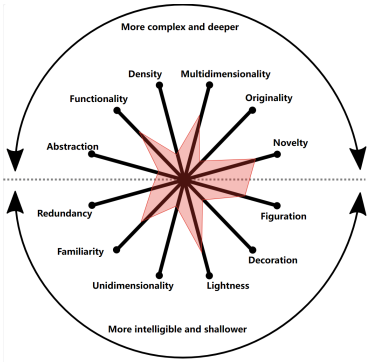

Fig. 1. A visualization wheel for each chart focus on increasing lightness, functionality, and multidimensionality.

5 Demonstration Scenarios

Our dashboard, designed with an array of modern web technologies including Vue.js and D3.js, serves as an interactive platform for analyzing California wildfires. Through the integration of libraries like vue-leaflet for mapping and element-plus for UI components, the dashboard offers a user-friendly interface and dynamic visualizations.

As Figure 2 shown, Users can explore temporal wildfire trends through a color-coded calendar. As Figure 3 shown, a dynamic scatter plot illustrates the number of fires and burned size over the years. As Figure 4 shown, We also use choropleth map to show the spatial distribution of wildfires, users can visually explore the distribution of fires. Selecting a county, it will show detailed statistics, enhancing the understanding of regional wildfire impact.This feature uses vue-leaflet to provide an interactive map that is both informative and intuitive. Given insights into wildfire causation, a pie chart breaks down the various causes of fires. Interactive elements allow users to click on segments of the pie chart for more detailed information.

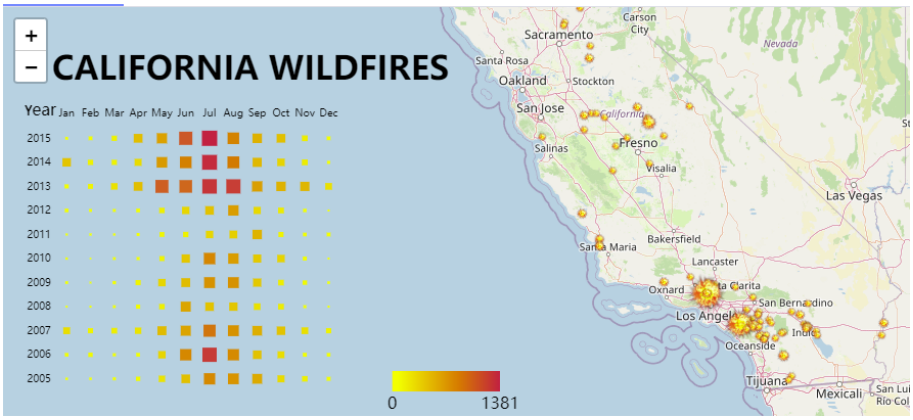

Fig. 2. Shows the peak fire seasons and temporal trends.

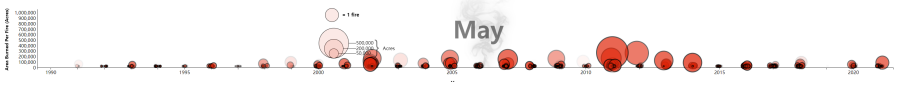

Fig. 3. Shows the number of fires and burned size over the years

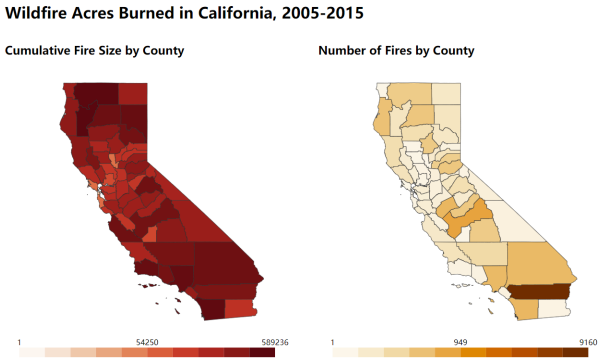

Fig. 4. Spatial distribution of wildfires

6 Conclusion

In conclusion, we have successfully utilized an extensive wildfire dataset to build an innovative interactive dashboard that provides a multidimensional analysis of wildfire occurrence. By analyzing spatial and temporal data from the Burned Area Boundaries dataset form MTBS and the 1.88 Million US Wildfires from Kaggle, we have provided a multifaceted view of wildfire dynamics in California. The dashboard employs tools like Vue.js, and D3.js in environmental data visualization and analysis, and create a user-friendly and informative experience.

Moving forward, we will look to enhance the predictive power of the dashboard by integrating more dynamic, real-time data sources. This will enable users to receive up-to-date information, which is critical for emergency response and resource allocation. In addition, we will incorporate feedback mechanisms to customize the dashboard's functionality to the specific needs of different user groups. By continually refining our tools and methodologies, we aspire to provide an even more robust platform that can adapt to the ever-changing landscape of wildfire management and actively contribute to disaster prevention and mitigation efforts.

References

1. Burned areas boundaries dataset. https://www.mtbs.gov/direct-download
2. Anna, F.: The growing danger of california wildfires. www.statista.com/chart/14462/california-wildfire-deadly (2022)
3. Chunyu, P., John T, K., Gregory S, T.W., Di, Y., Alex, G.M.: The season for large fires in southern california is projected to lengthen in a changing climate. Communication Earth Environment **3**(22), 3 (2022)
4. Tatman, R.: 1.88 million us wildfires. https://www.kaggle.com/datasets/rtatman/188-million-us-wildfires/data (2020)