California Wildfire Report

Graduation project | May 2025

---

1. Project Description

This report offers a detailed analysis of California wildfires from 2014 to early 2023. Using a combination of data preprocessing in Python, exploratory visualization through Tableau, and in-depth analysis of the original dataset, it reveals critical patterns in wildfire incidents, damages, and causes. The goal is to derive actionable insights to support strategic decision-making for disaster response, urban planning, and climate resilience.

---

2. Introduction & Objective

Introduction

Wildfires are one of California's most pressing environmental hazards, causing significant damage to homes, ecosystems, and human life. Understanding their trends, causes, and impacts is essential to mitigating future damage and strengthening community preparedness.

Objective

- Analyze wildfire incidents by year, season, month, and county.
- Examine causes and their relationships with destruction and injuries.
- Identify high-risk areas and contributing factors to damage.
- Recommend targeted prevention and response strategies.

---

3. Data Cleaning & Preprocessing

Initial Cleanup

- Removed duplicate rows and irrelevant columns.
- Standardized column names for consistency.
- Handled missing values with context-appropriate imputation (e.g., mode for categorical variables, mean/median for numeric).

Feature Engineering

- Extracted date parts (Year, Month, Season) from date fields.
- Created categorical groupings for causes and regions.
- Converted financial loss figures to millions for readability.

Data Transformation

- Encoded categorical variables for analytical use.
- Aggregated key metrics by time and geography (e.g., incidents per year/county).
- Ensured consistency in units (e.g., acres burned, millions of dollars lost).

Validation

- Ensured data integrity through summary statistics and visual inspection.
- Verified consistency with Tableau outputs.

---

4. Business Insights

A. Wildfire Impact Overview

- **Incidents:** 100 total wildfires from 2014–2023.
- **Damage:** $239.61 million in total financial losses.
- **Human Toll:** 4,000+ injuries, 100 fatalities, and 2.65 million acres burned.
- **Destruction:** ~94,000 homes destroyed, peaking in 2018–2021.

B. Trends Over Time

- **High-Impact Years:** The years 2018, 2021, and 2023 recorded the highest levels of wildfire destruction.
- **Monthly Patterns:** Incident frequency consistently peaks in October, reflecting a pronounced seasonal trend.
- **Seasonality:** Wildfire activity is concentrated during the **Fall and Summer** months, suggesting the need for intensified preparedness during these periods.

C. Cause Analysis

- **Human Activity** is the leading cause (38%), followed closely by **Unknown** and **Lightning** causes.
- All three cause types result in significant home destruction and injuries.

D. Geographic Analysis

- **Most Affected Counties:**
  - *Shasta*: 14% of total incidents
  - *Sonoma*, *Butte*, *San Diego*, *Mendocino* and *Riverside* also show high destruction
- **Injuries & Fatalities:** Highest in Mendocino, Shasta, and Napa Valley.

E. Economic Loss & Asset Damage

- **Asset Vulnerability:** Businesses and vehicles exhibit parallel damage trends, particularly in counties with high urban density, indicating a strong correlation between infrastructure concentration and economic exposure.
- **Primary Loss Driver:** Wildfires caused by **human activity** account for the highest financial damage—**over $85 million**, representing approximately **32.5%** of total recorded losses.

---

5. Recommendations

A. Focused Prevention Programs

- Target **Mendocino**, **Shasta**, and **Butte** counties with education campaigns and controlled burn strategies.
- Launch season-specific alerts and preparedness programs during **Fall** and **Summer**.

B. Cause Investigation & Mitigation

- Improve classification methods for fires listed as "Unknown."
- Strengthen policies addressing human-caused fires (campfires, electrical faults, arson).

C. Infrastructure & Community Resilience

- Invest in fire-resistant building materials in high-risk counties.
- Enhance evacuation planning and emergency response drills, especially in urban-wildland interface areas.

D. Smart Resource Deployment & Real-Time Monitoring

- Use predictive analytics and seasonal trends to proactively deploy firefighting and emergency teams in high-risk areas.
- Integrate real-time dashboards with live data feeds to monitor wildfire conditions, trigger automated alerts, and support rapid response decisions.
- Schedule regular data-driven strategy sessions to adjust resource allocation and enhance preparedness across counties.