# Analyzing the Impact of California Wildfires

### PART- IV [GROUP-06]

**Introduction and Background**

Wildfires are devastating in many ways. From burning hundreds of thousands of acres of land and destroying countless structures to causing extreme monetary damage, California wildfires cause nearly immeasurable damage in rural and urban areas across the state.

To better understand the damage caused by wildfires in California and increase awareness of the importance of wildfire prevention, it’s important to look at the numbers behind the worsening wildfires that the state has experienced over the past several years. This project aims to analyze the consequences of California wildfires by leveraging data on burned acreage, structural damage, financial losses, human casualties and air quality degradation. By exploring these factors, we hope to gain a deeper understanding of the growing wildfire crisis and its implications for communities across the state.

Additionally, we plan to extend our analysis to assess total disaster-related fatalities and financial costs at a national level across the United States. By exploring these factors, we hope to gain a deeper understanding of the growing wildfire crisis and its implications for communities across the state and the country. The motivation for this analysis stems from the alarming trend of worsening wildfires in California. Reports indicate that wildfires are becoming more destructive, with higher numbers of structures lost, increased financial damage, and more lives affected. News articles and research studies highlight the devastating impact of wildfires, such as the 2018 Camp Fire, which was the most destructive wildfire in California history, destroying nearly 18,800 structures and claiming 85 lives. The recent January 2025 wildfires, including the Palisades Fire, further illustrate the ongoing threat posed by wildfires. Research articles emphasize that wildfires not only cause immediate destruction but also lead to long-term consequences, such as deteriorating air quality, loss of biodiversity, and displacement of communities. Understanding these impacts is crucial for raising awareness and informing future policy decisions to mitigate the damage caused by wildfires.

Furthermore, by examining total disaster-related deaths and financial costs at a national scale, we aim to contextualize California’s wildfire crisis within the broader landscape of disaster management in the United States.

To further support this analysis, we incorporate a second dataset focused on air quality monitoring from 2000 to 2023 which is sourced from the U.S. Environmental Protection Agency (EPA) and provides daily records of key air pollutant values.

The air quality dataset is critical in understanding the secondary impacts of wildfires, particularly their effect on public health and environmental quality. By combining wildfire data with AQI data, this project can reveal patterns of pollutant spikes during major fire events and identify the regions most affected by poor air quality. This adds an important dimension to the analysis, showing that the consequences of wildfires extend far beyond burned structures and financial damage.

Together, both datasets provide a multifaceted view of the wildfire crisis, capturing not only the direct destruction but also the broader environmental and health-related outcomes. This combined approach allows for a more complete understanding of wildfire impacts and supports data-driven awareness, prevention, and policy-making efforts.

**Key Questions**

1. **1. Which areas in California experienced the most severe wildfire damage, and how were these areas affected in terms of air quality (CO AQI) during those incidents?**
2. **How does carbon monoxide pollution (CO AQI) fluctuate across wildfire-prone months in California?**
3. **Is there a measurable relationship between the average wildfire damage in a California county and the level of CO air pollution observed?**
4. **How have average wildfire damage and CO-related air pollution changed over time in California?**
5. **Which counties in California experienced the highest structural damage and air pollution from wildfires, and how do these impacts compare.**

**Datasets Description**

**Dataset 1 – California Wildfire Damage Data**

The dataset comes from Kaggle collected from the **California Department of Forestry and Fire Protection (**[**The California Wildfire Data🔥🔥🔥**](https://www.kaggle.com/datasets/vijayveersingh/the-california-wildfire-data)**).** The data set was created to track and analyze structural damage from wildfires for Fire prevention, Emergency response planning and Disaster Management.

The Data is collected from satellites and aerial platforms for evaluating the impact of wildfire structures. Likely to be funded by the state and federal fire agencies under wildfire prevention program. The data set is collected from the year **2013** to **2024** wildfire incidents and their damage. The Dataset contains **48 fields** and **100231 rows.** The dataset includes City, County specific to California.

**Key Variables**

**Incident Data:**

* Incident Name, Incident Number, Incident Start Date, Fire Name.
* Helps identify specific wildfires and their timelines.

**Location Data:**

* Latitude, Longitude, City, County, Zip Code, Community, CAL FIRE Unit
* Useful for mapping wildfire locations and understanding geographic patterns.

**Damage Data:**

* Damage Level (e.g, No Damage, Affected 1-9%)
* Key for assessing wildfire impact on structures.

**Structural Characteristics:**

* Structure Type, Structure Category, Roof Construction, Eaves, Vent Screen, Exterior Siding, Windowpane, Deck/Porch
* Help analyze which types of buildings are most vulnerable to wildfire.

**Property Value Data:**

* Assessed Improved Value, Year Built, APN (Parcel Number)
* Can be used to assess financial impacts and historical trends in construction.

**How will this dataset support our project goals?**

This dataset supports the project’s goals by providing valuable insights into the impact of wildfire on communities, helping to identify high-risk areas based on structural vulnerabilities. By mapping wildfire locations, we can observe trends in affected regions and assess patterns over time. Additionally, the dataset allows for an analysis of property values before and after wildfires, offering insights into the financial implications of these disasters.

**Limitations:**

Missing data in certain fields, such as structure distance to utilities, may affect the completeness of the analysis. Furthermore, since the dataset consists of historical records, it does not provide real-time data and cannot predict future wildfires. Lastly, its scope is primarily focused on structural damage rather than broader environmental or ecological impacts, which may limit a comprehensive understanding of wildfire consequences.

**Dataset 2 – California Wildfire Financial loss**

The Data is sourced from the **National Centre for Environmental Information [**[**Wildire\_data\_02**](https://www.ncei.noaa.gov/archive/archive-management-system/OAS/bin/prd/jquery/accession/download/209268)**]**. The purpose of the dataset was to track wildfire incidents, assess damage, and understand causes to improve future fire prevention and response strategies.

Covers period from 2020 to 2022 and the dataset contains 11 variables and 100 records.

**Key Variables:**

* **Incident\_ID** – Unique identifier for each wildfire event.
* **Date** – The date of the wildfire.
* **Location** – The county where the wildfire occurred.
* **Area\_Burned (Acres)** – The total land area affected by the fire.
* **Homes\_Destroyed** – Number of residential properties lost.
* **Businesses\_Destroyed** – Number of commercial properties damaged.
* **Vehicles\_Damaged** – Number of vehicles impacted.
* **Injuries** – Number of reported injuries.
* **Fatalities** – Number of deaths caused by the wildfire.
* **Estimated\_Financial\_Loss (Million $)** – Economic loss due to the wildfire.
* **Cause** – The reason behind the wildfire (e.g., Lightning, Human Activity, Unknown).

**How will this dataset support our project goals?**

The California Wildfire Damage dataset supports our project by providing insights into wildfire severity, causes, and economic impact, which we can analyze alongside weather conditions to identify contributing factors. With variables like area burned, property damage, injuries, fatalities, and financial loss, we can assess the destruction caused by wildfires and determine if certain causes (e.g., lightning, human activity) lead to more severe outcomes. By linking this dataset with weather data such as temperature, wind speed, and precipitation we can explore whether environmental conditions influence fire severity.

**Limitations:**

The dataset lacks direct meteorological data and is limited to 2020–2022, integrating it with weather datasets will help us uncover correlations and create meaningful visualizations to understand wildfire trends better.

**Dataset 3 – U.S. Pollution Data 2000 – 2024**

This dataset spans from January 2000 to 2023, comprising approximately 665,414 records and 21 columns. It focuses on analyzing key air pollutants across the United States such as Ozone (O3), Carbon Monoxide (CO), Sulphur Dioxide (SO2) and Nitrogen Dioxide (NO2)

Each entry captures daily pollutant measurements along with corresponding Air Quality Index (AQI) values, collected from official monitoring stations across the country. The data is particularly valuable for studying long-term trends and environmental health impacts, including those linked to wildfire events.

The dataset is primarily sourced from the U.S. Environmental Protection Agency (EPA) via the Air Quality System (AQS). It was further compiled and published on Kaggle by contributors BrendaSo and Angela Kim, who cleaned and structured the data for public use and research [[Kaggle](https://www.kaggle.com/datasets/guslovesmath/us-pollution-data-200-to-2022)].

**Purpose of the Dataset**

The EPA developed this dataset to Monitor and evaluate ambient air quality and to assess compliance with the Clean Air Act. Moreover to inform environmental policy and public health strategies. It is now used for research, data science projects, and public dashboards, like yours, to understand how events like wildfires impact air quality in affected regions.

**Key Variables**

* **Date** – The specific calendar date when the air quality measurements were taken (format: MM/DD/YY).
* **Address** – The full street address of the monitoring station that recorded the data.
* **State** – The U.S. state in which the monitoring station is located (e.g., Arizona).
* **County** – The county where the monitoring station is located (e.g., Maricopa).
* **City** – The city where the air quality was monitored.
* **O3 Mean** – The average concentration of ozone for that day, measured in parts per million (ppm).
* **O3 1st Max Value** – The highest 1-hour ozone concentration recorded.
* **O3 AQI** – The Air Quality Index value calculated for ozone, indicating how harmful the ozone level is to health.
* **CO Mean** – The daily average concentration of carbon monoxide in the air, measured in parts per million (ppm).
* **CO 1st Max Value** – The highest 1-hour carbon monoxide level recorded.
* **CO AQI** – The Air Quality Index for carbon monoxide.
* **SO2 Mean** – The average daily concentration of sulfur dioxide, measured in parts per billion (ppb).
* **SO2 1st Max Value** – The highest 1-hour sulfur dioxide concentration of the day.
* **SO2 AQI** – The calculated Air Quality Index score for sulfur dioxide.
* **NO2 Mean** – The daily average level of nitrogen dioxide, measured in parts per billion (ppb).
* **NO2 1st Max Value** – The highest 1-hour nitrogen dioxide level recorded that day.
* **NO2 AQI** – The Air Quality Index calculated for nitrogen dioxide.

**Geographic Scope**

While your visualizations may focus on California and specific counties, the full dataset includes measurements from multiple U.S. states, including cities like Phoenix, AZ, as shown in your preview.

**How will this dataset support our project goals?**

This dataset also helped us to measure air quality degradation during wildfire events and to draw comparisons between counties based on AQI levels. Moreover, to create temporal analyses (e.g., monthly/yearly AQI spikes) and to build compelling visuals that connect wildfire data to health-related outcomes.

**Limitations:**

Some regions may have incomplete data due to fewer monitoring stations. Does not directly include wildfire attribution must be cross-referenced with fire incident data. PM2.5 or smoke-specific metrics may be absent or underrepresented.

**Visualization Stories:**

**Wildfire + AQI Impact Map**

**A map of the state of california

AI-generated content may be incorrect.**

**Figure 1.1**

This choropleth map visualizes the average carbon monoxide (CO) Air Quality Index (AQI) values across counties in California during wildfire events, mapped geographically to uncover regional variations in air pollution levels linked to wildfire severity. Each county is shaded based on its average CO AQI during recorded wildfire incidents, with darker hues representing higher pollution levels.

**Datasets & Variables:**

**Dataset 1: California Wildfire Damage Data**

* County (Category): Geographic region where the wildfire incident occurred.
* Damage Level (Damage) (Category): Structural impact classification from “No Damage” to “Destroyed.” This field categorizes how severe wildfire affects buildings in a county.
* Latitude/Longitude (Numeric): Used for plotting exact location, when available.

**Dataset 3: U.S. Pollution Data 2000–2024 (pollution.csv)**

CO AQI (Numeric, Index): Air Quality Index for carbon monoxide

**Key values observed**

* Fresno County tops the chart with 6.0 AQI, followed by Kern County (5.0) and Sacramento County (4.9) suggesting hotspots of both wildfire activity and resulting air quality degradation.
* Several counties show 0.0 AQI, indicating either truly clean air or data unavailability, especially in less monitored or rural regions.
* Central Valley counties (e.g., Fresno, Kern, Sacramento) exhibit higher pollution, consistent with their topographical bowl-like structure, which can trap smoke and airborne pollutants.
* In contrast, Southern California counties such as San Diego (2.8) and Imperial (2.2) show moderate values, hinting at better atmospheric dispersion possibly due to coastal airflow or lower fire density in those zones.

**Limitations of Visualization**

* Temporal Resolution Limitation: Since the wildfire and pollution datasets use different date granularities, the AQI values represent approximate pollution conditions (+/- 2 days) from each wildfire event. Thus, there may be a mismatch in exact timing between pollution spikes and fire occurrences.
* Spatial Resolution Limitation: The data is aggregated at the county level, which does not reflect within-county variability. Pollution impact near dense population centers or fire origins may be lost due to averaging over a larger area.
* Zero Values Caution: Counties with 0.0 AQI might not reflect real clean air, but rather a lack of monitoring data or AQI reporting gaps.

**How This Chart Answers the Key Questions**

This map directly addresses the following key questions from your study:

**1. Which areas in California experienced the most severe wildfire damage, and how were these areas affected in terms of air quality (CO AQI)?**

* The map visually correlates air quality degradation with known wildfire-prone counties.
* Fresno, Kern, and Sacramento, all heavily fire-impacted regions, also reflect elevated CO AQI levels, showing that areas with greater wildfire activity do experience noticeable environmental consequences.
* It allows for pattern recognition: regions with repeated high structural damage and pollution can be flagged as dual-risk zones.

**2. Is there a measurable relationship between wildfire damage and CO air pollution?**

* Although not a scatter plot, this map lays the groundwork by showcasing counties with simultaneously high AQI and known wildfire severity.
* The visual contrast helps researchers identify where both pollution and structural risk coincide, motivating further statistical validation (which can be done in subsequent visuals or models).

**Chart Utility & Policy Implication**

This map is instrumental for:

* Disaster prepared teams to target air quality monitoring and health interventions in high-AQI counties post-wildfire.
* Urban planners and legislators reassess structural building codes in wildfire-prone zones that also show poor air quality outcomes.
* Environmental researchers seeking spatial AQI patterns linked to wildfire geography, topography, and wind behavior.

**Structure Type vs Average Damage Level**

A graph of a number of people

AI-generated content may be incorrect.

**Figure 2.1**

This horizontal bar chart presents the average wildfire damage scores for different types of structures across California. The damage score, ranging from 0 (No Damage) to 4 (Destroyed), is used to evaluate how severely each structure type is affected during wildfire incidents.

**Datasets & Variables:**

**Dataset1:** California Wildfire Damage Data

**Structure Type** *(Category)*: Classifies buildings like Residential, Mobile Home, Commercial, School, etc.

* **Damage Score** *(Ordinal – 0 to 4)*: Numeric mapping of wildfire damage level:
  + No Damage = 0
  + Affected (1-9%) = 1
  + Minor (10–25%) = 2
  + Major (26–50%) = 3
  + Destroyed (>50%) = 4

**Key observations from the chart:**

* Agricultural structures show the highest vulnerability with an average damage score of 4.0, indicating they are most frequently destroyed.
* Mobile homes, particularly Triple Wide (3.52) and Double Wide (3.31), follow closely behind, confirming their high susceptibility likely due to flammable materials and lack of foundation anchoring.
* Single Family Residences – Single Story (2.56) are more vulnerable than their multi-story counterparts (1.80), possibly due to exposure area and lack of elevation.
* At the other end of the spectrum, Hospitals (0.37) and Infrastructure (0.90) exhibit the lowest average damage scores, suggesting stronger building codes, fire-resistant design, and proximity to emergency response services.

**Limitations of Visualization**

* Small sample bias: Some structure types (e.g., Churches, Schools) may have fewer entries, which can distort their average score.
* Subjective scoring: The damage classification is based on post-incident reports, which may differ by evaluator or data collector.
* No temporal correlation: The chart does not distinguish whether high-damage incidents occurred in peak fire years or consistently across time.

**How This Chart Answers the Key Questions**

This chart directly supports:

**1. Which areas in California experienced the most severe wildfire damage, and how were these areas affected in terms of air quality (CO AQI)?**

* While this chart does not map geographic areas, it adds dimension to the “severity” aspect of the question by categorizing which structure types bear the brunt of destruction during wildfires.
* When combined with the AQI maps and regional financial loss visuals, this plot enriches the multi-faceted understanding of wildfire impact not just where, but what kind of structures are most at risk.

**2. Is there a measurable relationship between wildfire damage and exposure vulnerability?**

* Yes. The chart indicates a clear stratification of risk: mobile homes, utility sheds, and agriculture are at the top, implying the need for targeted protection efforts in such structure categories regardless of location.
* This insight can guide future policy and infrastructure decisions, such as offering incentives for retrofitting mobile homes or zoning restrictions for high-risk structures in fire-prone zones.

**Chart Utility & Policy Implication**

This visualization is particularly valuable for:

* Urban planners and engineers, who can use the data to revise fire-resilience standards for vulnerable buildings.
* Emergency managers, who now know which structures to prioritize during evacuation alerts or Fireline defenses.
* Insurance companies, who can calibrate premium models based on structure vulnerability.
* Communities, who may push for upgraded codes and better public awareness around fire-proof construction.

**CO Air Quality Index During Fire Season**

A graph of a graph

Description automatically generated with medium confidence

**Figure 1.3**

This box-and-whisker plot visualizes Carbon Monoxide (CO) AQI levels across different fire-season months in California. The months are grouped by fire relevance:

* May (Pre-Fire)
* June & July (Early Fire Season)
* August (Peak Fire Season)
* September (post-fire)
* Off Season (Remaining months)

Each group displays the distribution of daily CO AQI readings, highlighting median values, interquartile ranges, and outliers.

**Datasets & Variables Used:**

* **Dataset Used:** Dataset 3 – U.S. Pollution Data 2000–2024
* **CO AQI (Not Null)** (Numeric): Carbon Monoxide Air Quality Index, cleaned with IFNULL ([CO AQI], 0)
* **Pollution Date** (Date): Used to group months into wildfire-related categories
* **Fire Season** (Calculated Field):

1. CASE MONTH ([Pollution Date])
2. WHEN 5 THEN "May (Pre-Fire)"
3. WHEN 6 THEN "Jun"
4. WHEN 7 THEN "Jul"
5. WHEN 8 THEN "Aug"
6. WHEN 9 THEN "Sept (Post-Fire)"
7. ELSE "Off Season"
8. END

* **County** (Category): Used as a detail dimension

**Key observations:**

* August stands out with a notably higher median AQI, a wider IQR, and multiple extreme outliers (including a point >28,000 AQI), indicating intensified wildfire smoke and atmospheric degradation.
* May and September exhibit moderate CO AQI, suggesting the build-up and dissipation periods of wildfire impact.
* June and July maintain relatively lower medians, but some scatter points show early fire effects.
* Off-season months report consistently low CO AQI, as expected, due to reduced wildfire frequency.

**Limitations of Visualization**

* Indirect causation: CO AQI levels may result from traffic or industrial emissions, not just wildfires.
* Not tied to specific fire dates: These trends show seasonal influence, but don’t track pollutant spikes per wildfire.
* Outliers like >25,000 AQI are suspicious and should be verified for possible data entry errors or rare anomalies.

**How This Chart Answers the Key Question**

**3. How does carbon monoxide pollution (CO AQI) fluctuate across wildfire-prone months in California?**

* The chart clearly illustrates seasonal patterns, with peak pollution in August, aligning with known wildfire season intensity.
* It confirms that wildfire-prone months correlate with significant air quality degradation, especially in late summer.
* Health agencies and emergency planners can use this pattern to prepare for heightened respiratory risk windows, especially during August fire peaks.

**Chart Utility & Policy Implication**

This visual is valuable for:

* Public health departments planning early alerts or issuing masks during late summer months.
* Environmental agencies seeking evidence of wildfire-pollution seasonality.
* Wildfire mitigation teams timing preventive burns or response readiness based on expected AQI spikes.
* Researchers validating temporal air quality deterioration trends against meteorological and fire ignition data.

**Damage vs Air Pollution Over Time (Dual Line Chart)**

A graph of a graph showing the difference between air pollution and air pollution

Description automatically generated

**Figure 1.4**

This dual-axis line chart compares average wildfire structural damage (left y-axis, in blue) and average carbon monoxide (CO) Air Quality Index (AQI) (right y-axis, in red) over the years 2016 to 2024. The goal is to understand how physical destruction and environmental pollution co-evolve across fire seasons in California.

**Datasets & Variables Used:**

* **Year** (Both Datasets): Numeric field extracted from both Incident Date (Dataset 1) and Pollution Date (Dataset 3) for time-series analysis.
* **Avg Damage Score** (Dataset 1 – Extract): Numeric (0–4) field converted from categorical damage levels.
  + - * + 0 = No Damage
        + 4 = Destroyed
* **Avg CO AQI** (Dataset 3 – pollution.csv): Numeric (0–500 AQI) field representing the average carbon monoxide air quality index.

**Key insights:**

* Structural damage peaked in 2016 at 3.85, showing consistently high values until 2018, followed by a notable decline from 2019 onward.
* CO AQI spikes dramatically in 2019 (2.898), aligning with historical events like the Kincade Fire and Saddle Ridge Fire, confirming real-world correlation.
* Post-2019, CO AQI shows a steep drop, reaching zero by 2023–2024, which may point to:
  + Improved wildfire suppression or urban containment
  + Rural fires with less dense combustible materials
  + Or possibly missing or underreported data from sensors.
* From 2020 to 2024, while AQI remains low, structural damage persists around 1.0, suggesting that modern fires may be less polluting but still destructive to property.

**Limitations of Visualization**

* AQI = 0 in 2023 and 2024 could be due to missing pollution records, not necessarily clean air.
* This chart reflects statewide averages, which can flatten peaks or dilute local spikes in both damage and pollution.
* Does not control population density, fire size, or proximity to sensors, all of which influence CO AQI measurements.

**How This Chart Answers the Key Question**

**5. How does the average wildfire damage and CO-related air pollution change over time in California?**

* The chart clearly illustrates diverging trends: both damage and AQI were high pre-2020, but while structural losses declined moderately, pollution dropped sharply post-2019.
* These shifts highlight changes in fire patterns, possibly moving from urban-wildland interfaces to rural terrains.
* It allows for identification of critical years (e.g., 2019) where both damage and pollution peaked important for disaster response retrospectives.

**Chart Utility & Policy Implication**

This time-series visualization provides:

* A historical view of fire and pollution impact, essential for trend forecasting.
* Policy insight: shows the effectiveness (or gap) in air quality monitoring and disaster mitigation efforts over time.
* A tool for climate scientists, public health officials, and insurance firms to anticipate risk zones based on multi-year dynamics.

**Structural Damage vs CO AQI (Scatter Plot)**

A graph with numbers and points

Description automatically generated

**Figure 1.5**

This scatter plot depicts the relationship between average structural damage scores from wildfires and the average CO Air Quality Index (AQI) across various counties in California. Each data point represents a single county, plotted with:

* X-axis: Average wildfire damage score (0 = No Damage, 4 = Destroyed)
* Y-axis: Average CO AQI (Carbon Monoxide Index)

A trendline is included to assess the correlation between damage and air pollution.

**Variables & Datasets Used:**

* Damage Score (Dataset 1): Numeric field from 0 (No Damage) to 4 (Destroyed)
* CO AQI (No Null) (Dataset 3): Average carbon monoxide AQI per county
* County: Used to label each data point

**Key insights from the plot**:

* The overall trendline is slightly sloping downward, suggesting a weak inverse relationship as damage increases, AQI tends to slightly decrease.
* Alameda, Santa Clara, and Solano counties report high CO AQI values despite moderate to low structural damage, which could indicate smoke drift from fires elsewhere or high urban pollution unrelated to fires.
* In contrast, Butte, Shasta, Mono, and Kings counties have high structural damage scores but relatively low AQI, which could stem from rural fires with less industrial emissions or rapid pollutant dispersion.
* Several counties (e.g., Colusa, Alpine, Glenn) have low scores on both axes, confirming they are less affected structurally and environmentally.
* The spread of points highlights significant inter-county variability, implying that wildfire severity doesn’t always predict pollution levels due to local geography, wind patterns, or monitoring density.

**Limitations of Visualization**

* County-level aggregation masks local variation (e.g., urban centers vs forested outskirts).
* Smoke drift and external pollution sources aren’t captured counties may record high AQI from neighboring wildfires.
* Does not differentiate fire size or type, which can affect pollution but not necessarily cause structural loss.
* Some counties with extreme values might reflect monitoring or reporting differences, not actual environmental conditions.

**How This Chart Answers the Key Question**

**4. Is there a measurable relationship between the average wildfire damage in a California county and the level of CO air pollution observed?**

* The chart reveals that no strong correlation exists, but it’s not random for either certain county consistently falls into distinct behavior groups:
  + Urban-adjacent (high AQI, low damage): likely affected by drifting smoke and traffic-related emissions.
  + Rural-high fire zones (high damage, low AQI): more destruction but lower pollution retention.
* This insight is crucial: it shows you cannot assume high pollution simply from high damage the relationship is mediated by topography, population density, and wind flow.

**Chart Utility & Policy Implication**

This scatter plot is highly valuable for:

* Environmental planners need to understand why some high-damage areas escape heavy pollution is useful for designing targeted air monitoring stations.
* Health response teams identifying non-intuitive pollution hotspots, even in low-damage counties.
* Research studies focusing on cross-domain wildfire impact modeling this chart links the physical and environmental dimensions in one visualization.
* Policymakers who can’t use structural loss alone as a proxy for environmental danger both must be addressed in planning and response.

**County-wise Structural Damage and CO AQI Comparison**

A screenshot of a graph

Description automatically generated

**Figure1.6**

This tabular visualization compiles data across counties in California to compare wildfire impact intensity using multiple indicators:

* Average Structural Damage Score (0 to 4)
* Total Number of Structures Affected
* Maximum CO AQI (worst-case air pollution during fire events)
* Average CO AQI (chronic air pollution exposure)

The goal is to assess both physical destruction and environmental degradation caused by wildfires across different regions.

**Datasets & Fields Used:**

* **County** (Both Datasets):  
  Grouped geographic unit used for aggregating both structural damage and air quality metrics across datasets.
* **Avg. Damage Severity** (Dataset 1):  
  Aggregation: AVG (Damage Score)  
  Numeric value (0–4) converted from textual damage levels (e.g., “Minor”, “Destroyed”) to quantify average structural impact per county.
* **Total Structures Affected** (Dataset 1):  
  Aggregation: COUNT ([Structure Type])  
  Total number of building records affected or evaluated in each county.
* **Max. CO AQI** (Dataset 3):  
  Aggregation: MAX (CO AQI)  
  The highest carbon monoxide AQI reading observed in each county captures worst-case air quality during fire events.
* **Avg. CO AQI** (Dataset 3):  
  Aggregation: AVG (CO AQI)  
  The average air quality index value for carbon monoxide, reflecting chronic exposure during the reporting period.

**Key insights:**

* Los Angeles County tops the chart with 11,492 affected structures and a Max CO AQI of 16, reflecting the most extensive damage and severe pollution making it the epicenter of fire impact.
* Contra Costa and Solano Counties show high average AQI levels with lower structural damage, likely affected by smoke drift or industrial pollution, underscoring the environmental burden even in less structurally affected areas.
* Ventura, Santa Barbara, and San Diego represent balanced risk zones, with moderate damage and AQI, indicating consistent wildfire exposure both structurally and atmospherically.
* Fresno and Riverside Counties, though not topping AQI, show substantial structure damage, confirming vulnerability despite lower pollution averages.
* Counties with Max CO AQI ≥ 8 (e.g., Solano, Orange) highlight potential chronic health risks and warrant immediate public health and air monitoring interventions.

**Limitations of Visualization**

* Static table view may not reflect the dynamic timing of fires or pollution surges.
* AQI measures may vary due to sensor placement density, affecting comparability across counties.
* Some counties may lack complete data for either structure or pollution metrics.

**How This Chart Answers the Key Question**

**6. Which counties in California experienced the highest structural damage and air pollution from wildfires, and how do these impacts compare?**

* This table provides a direct, comparative analysis between counties, bridging infrastructure damage and air quality deterioration.
* Enables side-by-side comparison of:
  + Where damage is high, but air quality is not (e.g., Fresno)
  + Where pollution is high despite low structure loss (e.g., Solano, Contra Costa)
  + Where both are elevated (e.g., Los Angeles)
* Help stakeholders triangulate true high-risk zones, factoring in both human exposure and physical vulnerability.

**Chart Utility & Policy Implication**

This comprehensive table is a powerful tool for:

* Disaster response planning, allowing geographic prioritization of evacuation, mitigation, and air quality alerts.
* Urban planning teams, helping tailor building regulations and monitoring setups in high-risk counties.
* Public health officials, focusing air-quality campaigns and emergency protocols where pollution extremes exist despite fewer structural losses.
* State-wide resource allocation, directing funds and fire readiness equipment based on total impact scope not just damage count.

**Impact of Wildfires on Homes, Businesses, and Vehicles Over Time**

The stacked bar chart illustrates the impact of wildfires on homes, businesses, and vehicles from 2014 to 2023, revealing significant fluctuations in destruction levels over time. Peak destruction years such as 2014, 2018, and 2021 experienced the highest losses, with homes being the most affected, followed by businesses and vehicles. Visualization highlights the increasing frequency and severity of wildfires, particularly in recent years, emphasizing the urgent need for enhanced fire mitigation efforts. While some years, like 2017 and 2019, show relatively lower damage, the consistent rise in destruction from 2020 onward indicates that wildfires remain a growing threat. The dominance of home losses in each year suggests a critical need for improved residential fire protection strategies, including fire-resistant building materials and defensible space planning. This chart supports our goal of understanding wildfire trends, informing policy decisions, and strengthening emergency preparedness by identifying high-risk years and the disproportionate impact on residential areas.

**Variables Used:**

* Date and Properties damaged

**Dataset:** Wildfire Damage Dataset (Dataset 2)

A graph of different colored bars

AI-generated content may be incorrect.

**Figure 1.7**

**Financial Impact of Wildfires Across California County**

The financial loss bar chart presents the estimated monetary impact of wildfires across various counties, highlighting significant disparities in losses. Shasta County and Butte County exhibit the highest financial damages, surpassing $30 billion, indicating their vulnerability to severe wildfires. Mendocino, San Diego, and Sonoma Counties also demonstrate substantial losses, suggesting that wildfires in these regions have caused extensive property and infrastructure damage. Conversely, counties such as Santa Barbara and Los Angeles report lower financial losses, potentially due to better mitigation strategies or lower wildfire occurrences. This visualization effectively supports the goal of identifying high-risk areas where wildfire-related financial burdens are most severe, helping policymakers and emergency responders prioritize resource allocation and implement targeted prevention strategies.

**Variables Used:**

* Location and Estimated Financial Loss

**Dataset:** Wildfire Damage Dataset (Dataset 2)

**A graph with different colored bars

AI-generated content may be incorrect.**

**Figure 1.8**

**Wildfires Toll: Fatalities and Injuries by Cause and County**

Wildfires in California have not only caused widespread destruction to property but have also led to significant human casualties, as shown in the visualization of fatalities and injuries across various counties. The data highlights that counties such as Shasta, Butte, and Sonoma have experienced disproportionately high numbers of wildfire-related deaths and injuries, largely due to major fire events like the 2018 Camp Fire, which claimed over 85 lives. This visualization supports our research goal by illustrating the broader impact of wildfires beyond financial losses, emphasizing the urgent need for improved evacuation strategies and emergency response systems. By analyzing the distribution of casualties, we can identify high-risk areas where additional resources, such as early warning systems and community education programs, are needed to enhance preparedness. The data underscores the necessity of strengthening fire prevention policies and ensuring that residents in vulnerable regions have access to clear safety plans. Understanding the human toll of wildfires is crucial in driving policy decisions that prioritize both structural resilience and human safety in wildfire-prone areas.

**Variables Used:**

* Year, Fatalities, Injured and Cause.

**Dataset:** Wildfire Damage Dataset (Dataset 2)

A graph of different colored squares

AI-generated content may be incorrect.

**Figure 1.9**

**How Temperature Stability and Precipitation Fluctuations Influence Wildfire Risks**

This chart captures the relationship between precipitation fluctuations and the relatively stable trend of maximum temperature over time, providing key insights into climate patterns and their potential impact on wildfires. The precipitation trend shows significant variability, with periods of heavy rainfall followed by dry spells, which can contribute to wildfire risks by first promoting vegetation growth and then leaving dry fuel for fires during drought periods. In contrast, the maximum temperature remains relatively steady with minor fluctuations, suggesting that while extreme temperature spikes may not be the primary driver, sustained warmth over the years can still play a role in drying vegetation and extending fire seasons. This visualization helps address our project’s goal of understanding how climate variability influences wildfire frequency and severity. By examining these trends, we can assess how precipitation-driven fuel build up, combined with persistent warm conditions, increases fire risks. These findings reinforce the importance of integrating climate data into wildfire preparedness and mitigation strategies to better anticipate and manage future fire hazards.

**Variables Used:**

* Year, Avg Precipitation, Max temperature.

**Dataset:** California Weather Dataset (Dataset 3)

**A graph of a graph

AI-generated content may be incorrect.**

**Figure 10.0**

**Average CO Air Quality Index (AQI) by County**

This line chart compares the average carbon monoxide AQI across selected California counties over time. The chart reveals differing levels of exposure to wildfire-related air pollution. For instance, Los Angeles and Orange Counties experience some of the highest AQI levels in 2019–2020, reflecting intense wildfire smoke and urban vulnerability. In contrast, counties like San Diego and Contra Costa maintain more stable AQI values, potentially due to fewer fires or better air dispersion. This visualization supports the analysis of regional vulnerability to air quality degradation caused by wildfires and can help inform public health response and resource prioritization.

**Variables Used:** County (Location), Average CO AQI and Year of Date

**Dataset:** California Wildfire Damage Data (Dataset 1) and U.S Pollution Data (Dataset 3)

A graph of different colored lines

Description automatically generated

**Figure 11.1**

**Value of Property Damaged vs AQI Over Time**

This dual-axis line chart illustrates the relationship between the **average carbon monoxide (CO) Air Quality Index (AQI)** and the **total assessed property damage value** across wildfire-affected areas in California from 2014 to 2024. The green line shows trends in average CO AQI, while the red line indicates the total financial value of property damages. A sharp spike in damage value in 2018 aligns with the historically devastating Camp Fire, suggesting a direct connection between wildfire intensity and economic loss. Meanwhile, elevated AQI levels in 2018–2020 coincide with major fire seasons, reflecting how wildfire smoke contributes to worsening air quality. The overall downward trend in damage post-2020 may reflect fewer large-scale wildfires or more effective mitigation efforts. This chart supports the project's goal of connecting **air quality degradation and economic impact,** highlighting the **dual consequences** of wildfire activity.

**Variables Used:** Average CO AQI (Air Quality Index), Assessed Improved Value (in USD) and Year of Date

**Dataset:** California Wildfire Damage Data (Dataset 1) and U.S Pollution Data (Dataset 3)

A graph with red lines and numbers

Description automatically generated

**Figure 12.1**

**Dashboard**

**Conclusion**

The charts collectively demonstrate the devastating impact of wildfires on California’s communities. While some years show a decline in damage, major fire incidents continue to cause billions of dollars in destruction. Counties with repeated high losses require enhanced fire prevention measures, and early warning systems should be strengthened to minimize future damages. By analyzing these visual trends, policymakers and researchers can work towards a more fire-resilient future.

**References**

1. CNRA GIS Open Data, California Natural Resources Agency, [[online](https://gis.data.ca.gov/)].
2. Frontline Wildfire Defense, “California Wildfire Damage,” Wildfire News and Resources, [[Online](https://www.frontlinewildfire.com/wildfire-news-and-resources/california-wildfire-damage/)].
3. Internet Geography, “Wildfires in California: Causes, Effects and Responses,” [[Online](https://www.internetgeography.net/wildfires-in-california-causes-effects-and-responses/)].
4. National Interagency Fire Center (NIFC), “Fire Prevention, Education & Mitigation,” [[Online](https://www.nifc.gov/fire-information/fire-prevention-education-mitigation)].
5. J. Cheng, “Fire Behavior and Climate Relationships,” U.S. Forest Service Research Paper, GTR-NRS-P-84, 2009. [[Online](https://www.nrs.fs.usda.gov/pubs/gtr/gtr-nrs-p-84papers/04cheng-p-84.pdf)].
6. PRD OAS, “Download Accessions,” [Online]. Available through internal systems or custom access portal.
7. California Weather and Fire Prediction Dataset (1984–2025), With Engineered Features. [Online]. Source unspecified.
8. Kaggle. (2023). California Wildfire Dataset. [The California Wildfire Data🔥🔥🔥](https://www.kaggle.com/datasets/vijayveersingh/the-california-wildfire-data?resource=download).
9. Abatzoglou, J. T., Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. Proceedings of the National Academy of Sciences, 113(42), 11770-11775.
10. Kaggle. (2023). California Wildfire Dataset. [U.S Pollution 2020-2024](1.%09https:/www.kaggle.com/datasets/guslovesmath/us-pollution-data-200-to-2022)

**Contributions**

Darsini Laskmiah  
**Contributions:**

* Responsible for developing the map and bar chart visualizations in Tableau.
* Assisted in data cleaning and preprocessing for better visualization quality.
* Wrote the Introduction and Background sections of the proposal.

Hari Prasannaa Thangavel Ravi  
**Contributions:**

* Led the research and selection of datasets for the project.
* Conducted initial analysis of the data and identified key variables to focus on.
* Responsible for developing the line and heatmap visualizations in Tableau.

Abirham Getie  
**Contributions:**

* Contributed to the writing of the Objectives and Goals section.
* Created preliminary charts in Tableau to test dataset compatibility.