

Project 3 –

An Analysis of California Wildfires

Natalie Ericson
Peter Solis
Ralph Gaston
Timothy Coleman

Table of contents

01

Data Collection/Cleaning

02

Creating the Code

03

The Visuals

04

Data Analysis and Findings



Driving Questions

**Does more precipitation in the rainy season
lead to a worse fire season in California?**



Driving Questions

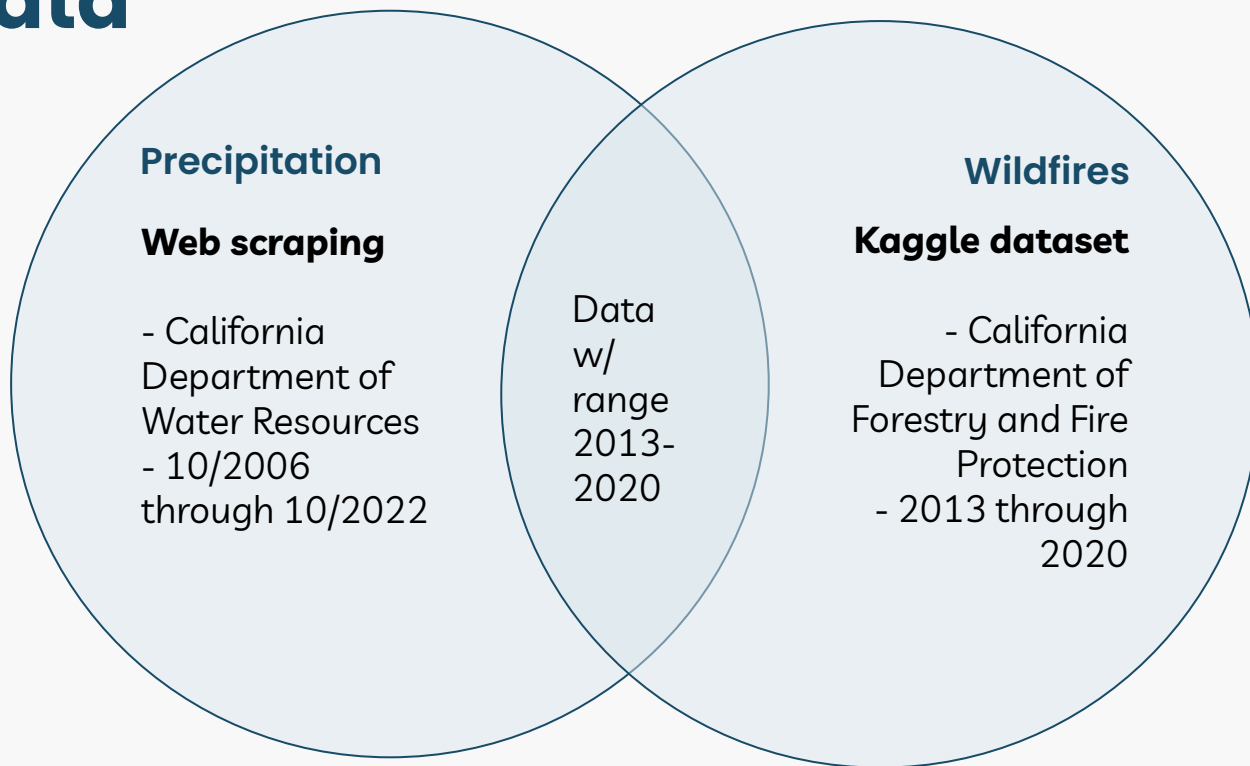
**Using data from past years (2013–2020),
can we generalize wildfire activity
trends in California for upcoming years?**



01

Data Collection & Cleaning

The Data





Web Scraping

- California Data Exchange Center provides statistics on precipitation from multiple weather stations
- Primary table gives monthly info including:
 - Inches of precipitation
 - Average precipitation for that month
 - Station ID
- Scrape secondary table for station info:
 - Full station name
 - Latitude, Longitude

Data Cleaning

- Downloaded California_Fire_Incidents.csv
- Read this in as wildfires_df to manipulate with Pandas
- Narrowed down the columns to the ones of interest and renamed them
- Converted the 'Start_Date' column to string
 - Then used .to_datetime function with Y-M-D format
 - This enabled us to use dt.Month and dt.Year functions to create new 'Month' and 'Year' columns for each fire
 - Dropped initial 'Start_Date' column
- Dropped any rows that did not have a Latitude or Longitude value
- Wrote final wildfires_clean_df to final wildfires.csv for use

The Clean Data

rainfall.csv

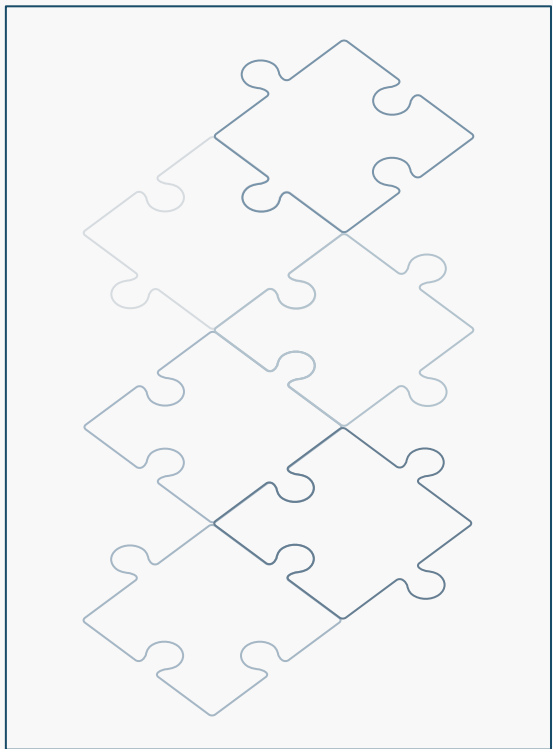
Station_ID	BIGINT
Index_Num	TEXT
Pcpn_In	FLOAT
Avg_Pcpn	FLOAT
Avg_Pcpn_Percent	BIGINT
Month	BIGINT
Year	BIGINT

rain_stations.csv

Station_ID	TEXT
Station_Name	TEXT
Longitude	FLOAT
Latitude	FLOAT
Elevation	BIGINT
More_Info	TEXT

wildfires.csv

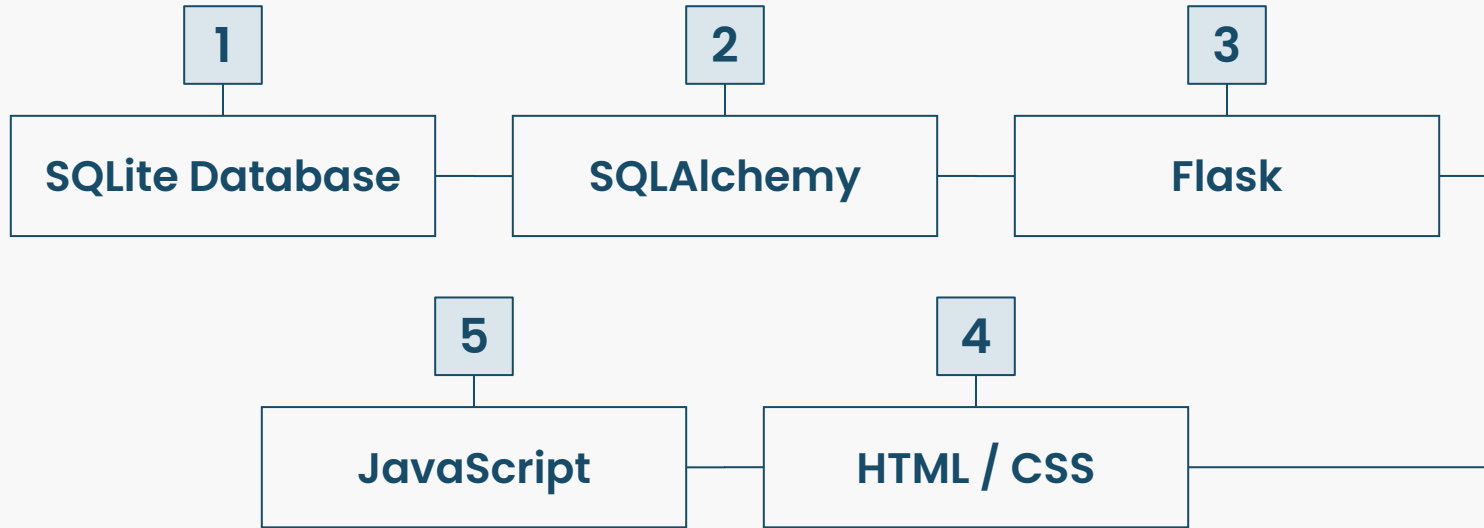
Unique_ID	TEXT
Acres_Burned	FLOAT
Cal_Fire_Incident	BOOLEAN
Latitude	FLOAT
Longitude	FLOAT
Month	BIGINT
Year	BIGINT



02

Creating the Code

Development Tools



Leaflet.Sync

A Leaflet library that creates a synchronized view of two maps

- Allows us to display markers for precipitation data and wildfire data side-by-side without confusion with overlapping
- Can change the base layer to toggle between street, satellite, and topographical views separately
- Zoom and scroll features are synced so view stays consistent across both maps





03

The Visuals



Behind the Visuals

Available Routes

`/years`

All years that contain both fire and precipitation data

`/map_markers/yyyy`

Data for the map displaying fire locations and precip. amounts

`/rain_bar/yyyy`

Data for the bar chart with monthly precipitation averages

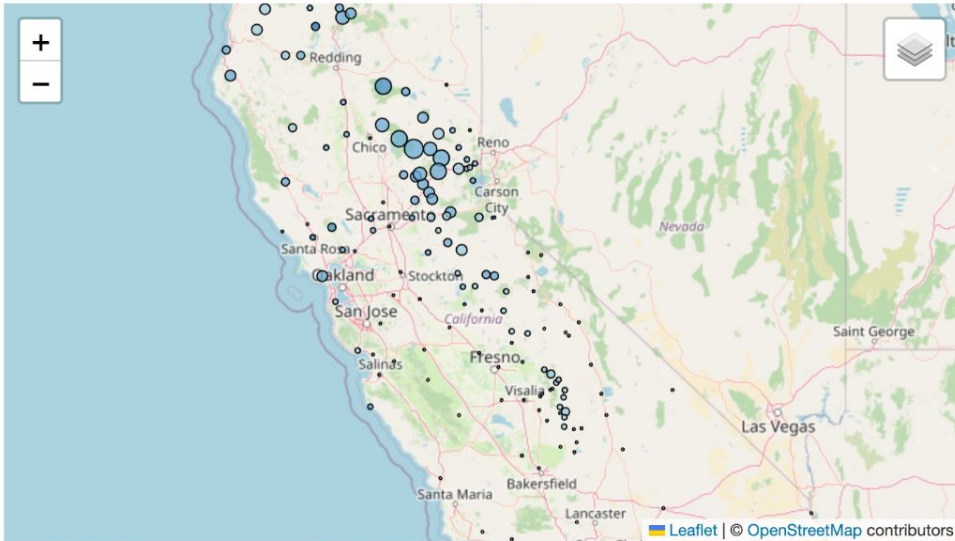
`/fire_bar/yyyy`

Data for the bar chart with monthly fire count and acres burned

`/month_line`

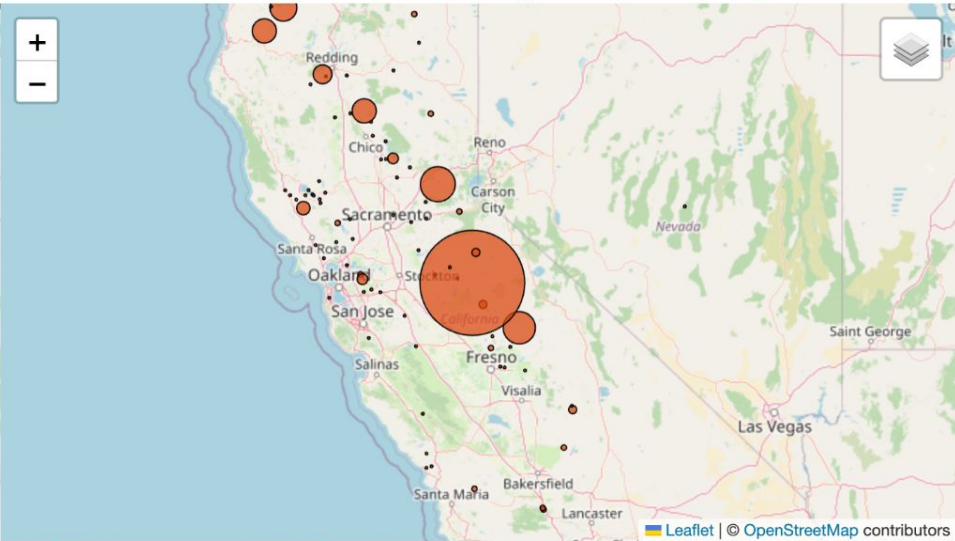
Data for the line chart comparing precipitation and fire data

2013 Rainy Season Data (November - June)



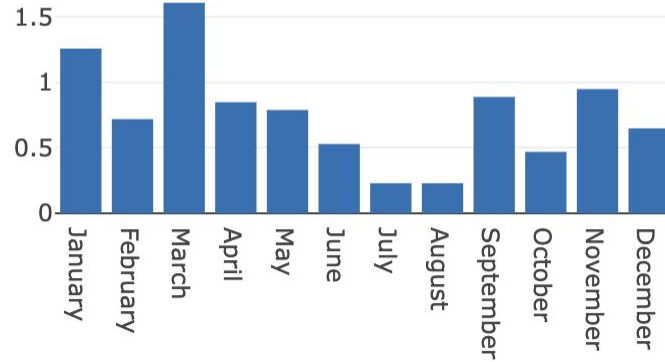
- Circle size represents amount of precipitation during the rainy season (bigger = more rain).
- Darkness of circle represents percent of average precipitation for the season (darker = more rain than usual).
- [Data Source](#)
- Note: Density of circles just shows density of stations where data was measured, so low density does not mean low precipitation.
- Click a marker for more info.

2013 Fire Season Data (June - November)

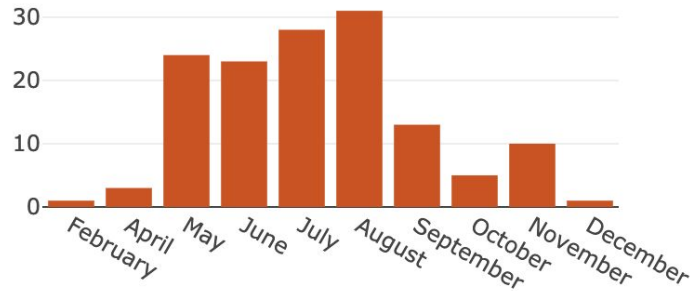


- Circle size represents number of acres burned.
- [Dataset](#)
- [Data Source](#)
- Click a marker for more info.

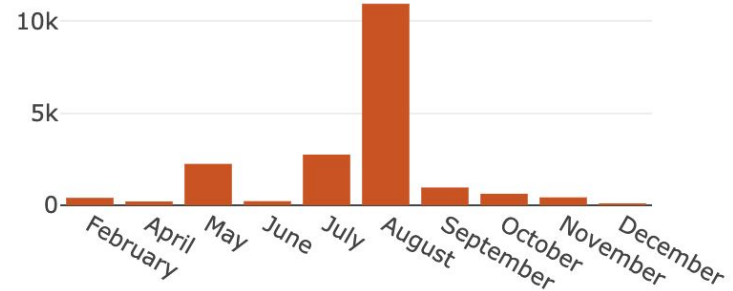
Monthly Avg Precipitation in CA - 2013



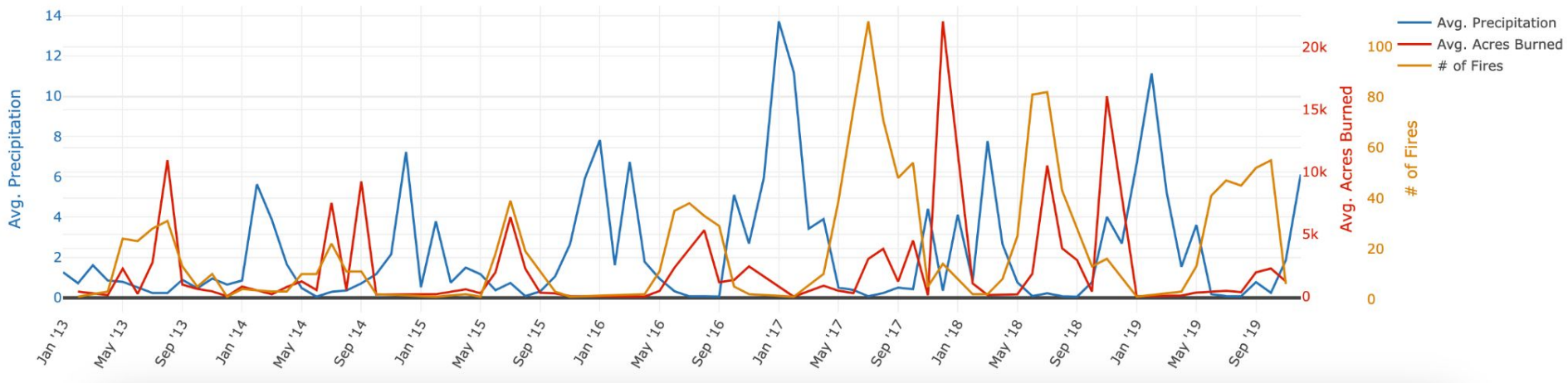
Monthly # of Fires in CA - 2013

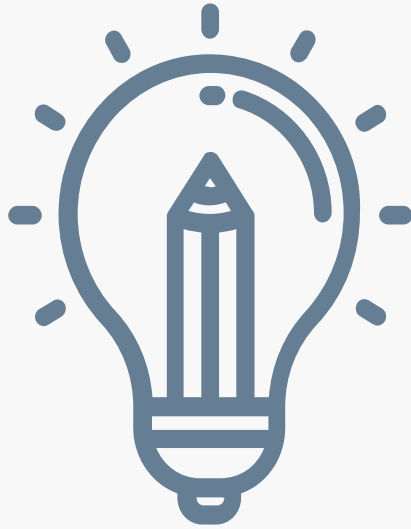


Monthly Avg Acres Burned in CA - 2013



Monthly Precipitation and Wildfire Trends (2013-2019)





04

Data Analysis and Findings



Driving Questions

**Does more precipitation in the rainy season
lead to a worse fire season in California?**

Conclusions

- There is not enough obvious data to define a specific trend between precipitation and wildfires
 - The amount of precipitation in the rainy season *does appear to have a **positive** correlation* with the number of fires during the peak fire season
 - The amount of precipitation in the rainy season *does **not** appear to have a strong correlation* with the number of acres burned in a fire season
 - The number of acres burned could be due to factors other than precipitation
- Why might this be the case?
 - More rain leads to more undergrowth
 - This may make it easier for fires to start



Driving Questions

**Using data from past years (2013–2020),
can we generalize wildfire activity
trends in California for upcoming years?**



Driving Questions

**Using data from past years (2013–2020),
can we generalize wildfire activity
trends in California for upcoming years?**

... probably not

Limitations

- Not enough time to run statistical analysis on the data to get more concrete answers, such as a calculating a correlation coefficient
- Relatively small dataset for the type of information we're looking at
 - Could include additional data on:
 - drought seasons or intense drought events
 - fire department response times
 - ease of access to the location of fire
 - population of the area
 - other environmental factors



Future Possibilities

- Can we increase the number of years of available data to provide more insight into long term patterns?
- Can we examine fire department response times in correlation with our data on number of acres burned to draw additional conclusions on the risk a potential fire poses to a community?
- Can we examine the cause of a fire in correlation with its location to tell us if there's a relationship between campsite occupation and wildfire occurrence?





Thanks!

Any questions?

CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)

Please keep this slide for attribution
