Exploring Wildfires: Can Only *We* Prevent Forest Fires?

Gopalika Sharma and Surya Menon

Background

- Climate change has resulted in lengthening forest fire seasons
- In 2020 California has had:
  - **8,685** fire instances
  - Over 4 million acres burned
  - 9,247 damaged structures
  - At least 31 fatalities

2020 California Fires Map

Problem Statement

Research to help develop strategies to manage the impact of future fires:

- Spatial analysis of California wildfires
- Prediction of fire features
  - Fire Size
  - Fire Cause

Main Datasets:

- Historic US wildfires (1992-2015)
- Portugal forest fire incidents

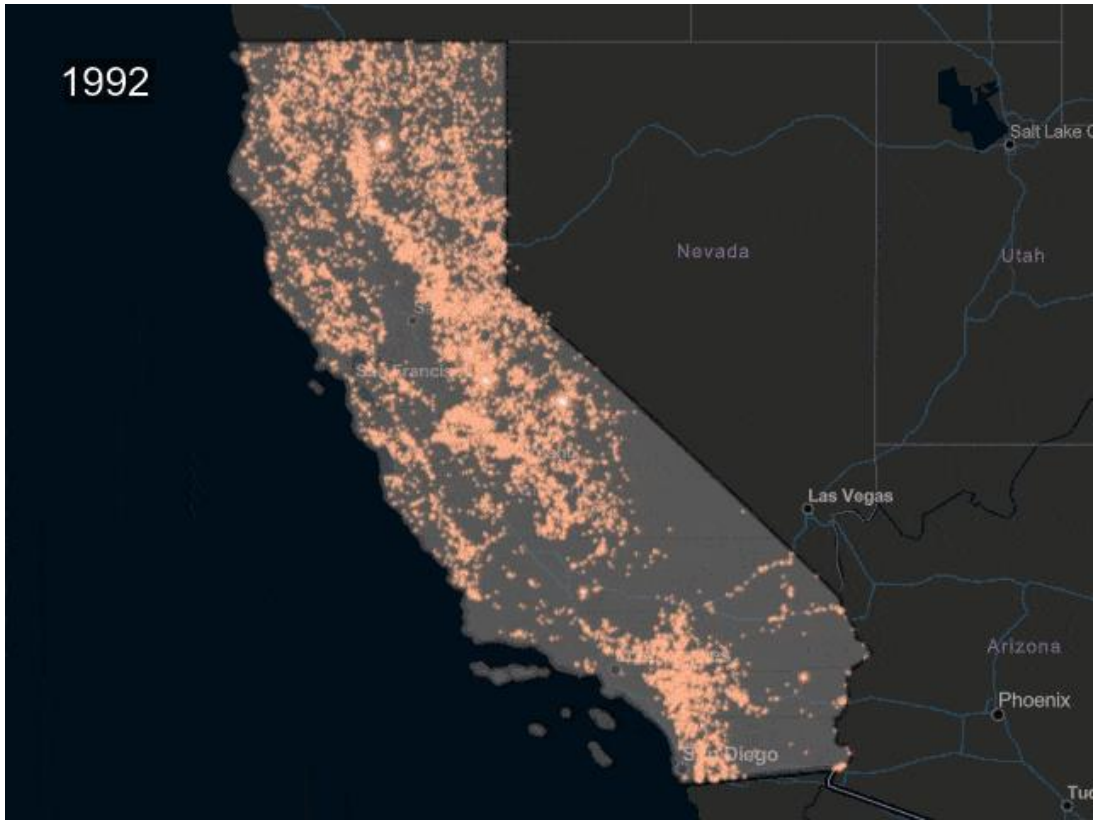
1.

Spatial Analysis

Spatial Analysis - Methods

- ArcGIS to observe patterns and distribution of fires
  - Supplemental spatial datasets
- Proximity Analysis
- Emerging Hot Spot Analysis
- California Wildfires Dashboard
  - kepler.gl
  - plotly Dash

California fires have increased in size and severity over time

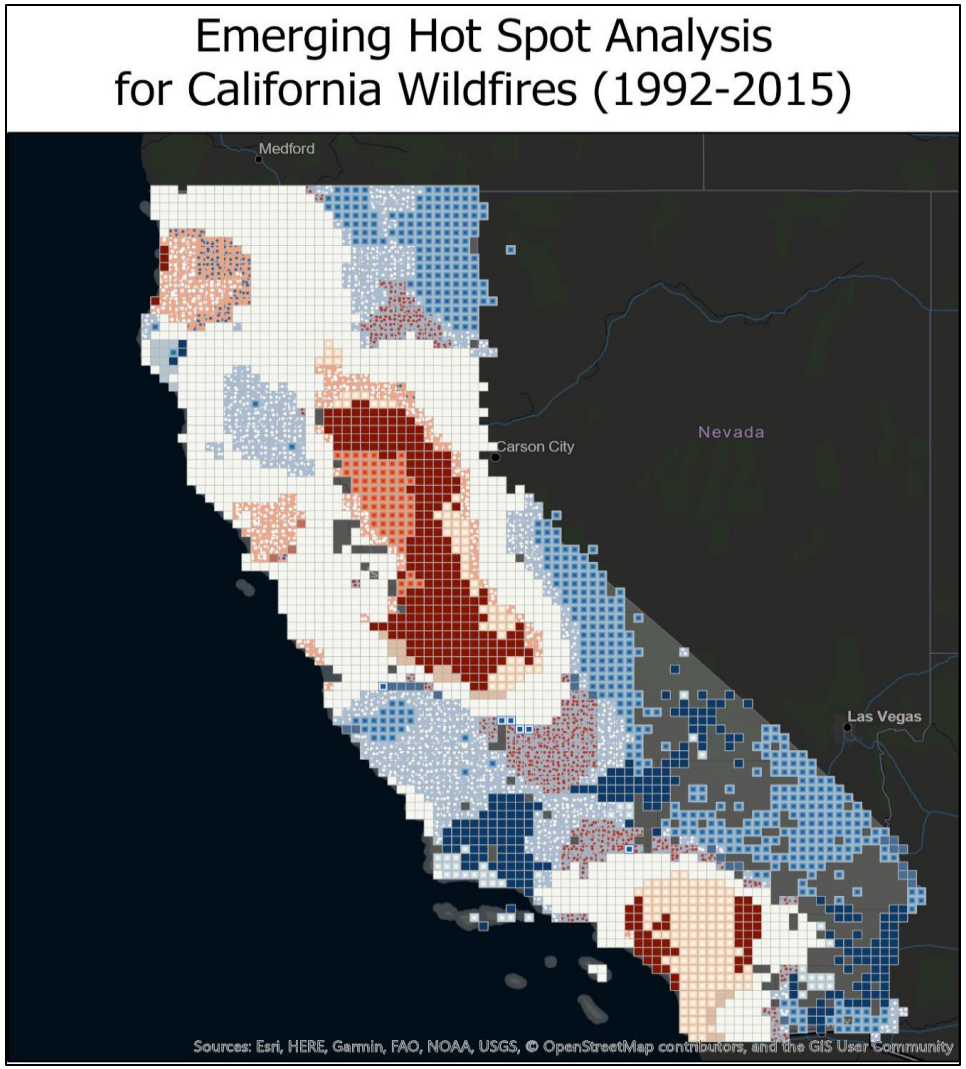


Time Lapse of California Wildfires (1992-2015)

Persistent fires in Sacramento Valley area

Hot Spot Designation

- New Hot Spot
- Consecutive Hot Spot
- Intensifying Hot Spot
- Persistent Hot Spot
- Diminishing Hot Spot
- Sporadic Hot Spot
- Oscillating Hot Spot
- Historical Hot Spot
- New Cold Spot
- Consecutive Cold Spot
- Intensifying Cold Spot
- Persistent Cold Spot
- Diminishing Cold Spot
- Sporadic Cold Spot
- Oscillating Cold Spot
- Historical Cold Spot
- No Pattern Detected

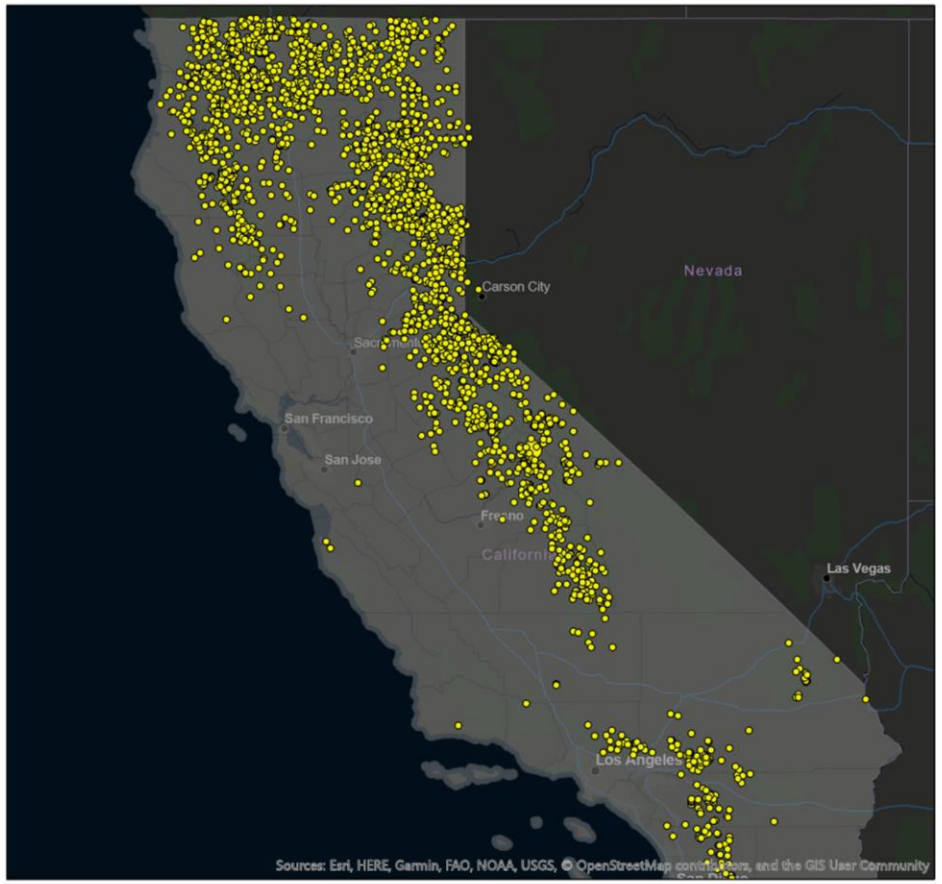


Increasing occurrences of climate-related wildfires

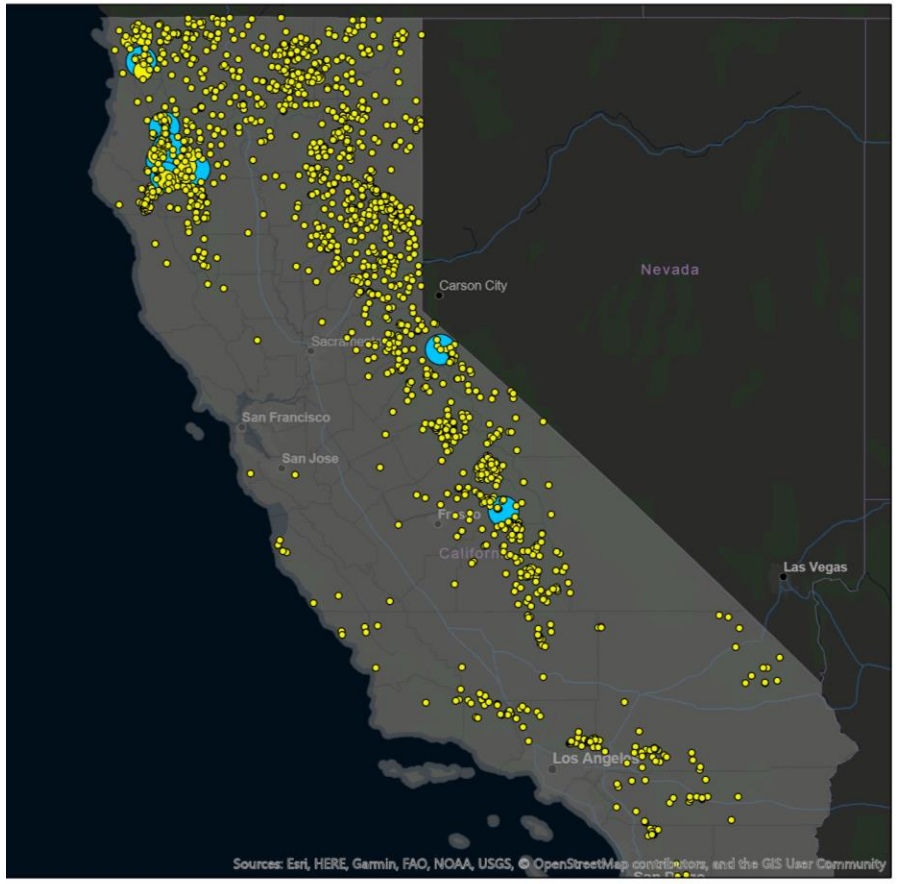
Fire Size (Acres)

- ≤ 2626.747252
- ≤ 7251.99534
- ≤ 11877.243428
- ≤ 151623.0

1992 CA Lightning Fires by Size (Acres)



2015 CA Lightning Fires by Size (Acres)



Hospitals could be better distributed in high count zones

CA Fire Stations

- CA Fire Stations

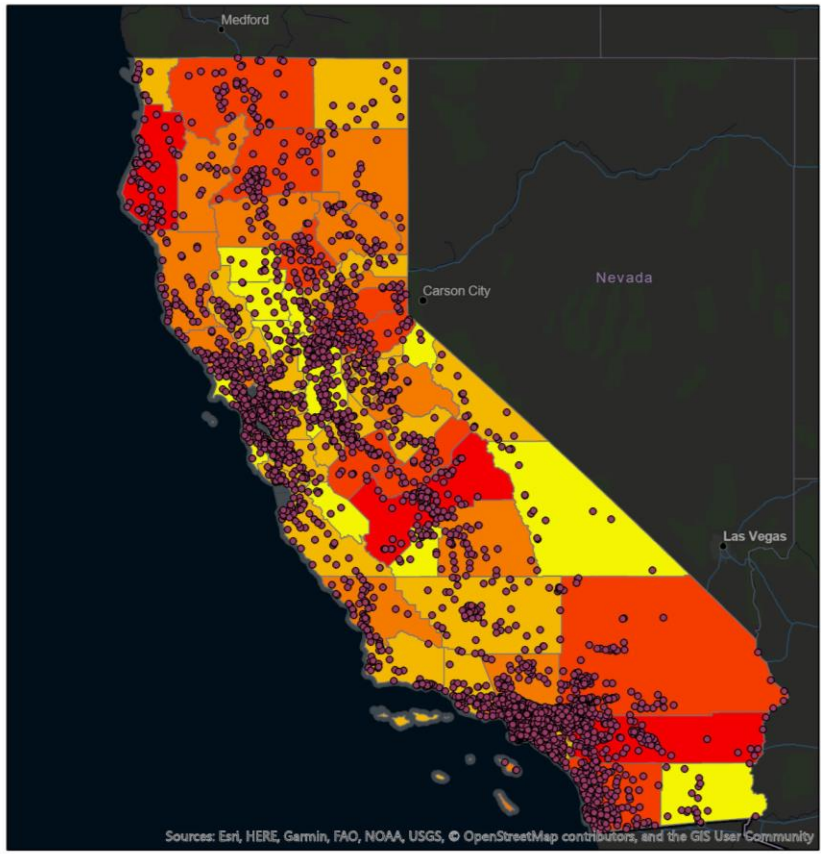
CA Healthcare Facilities

- CA Healthcare Facilities

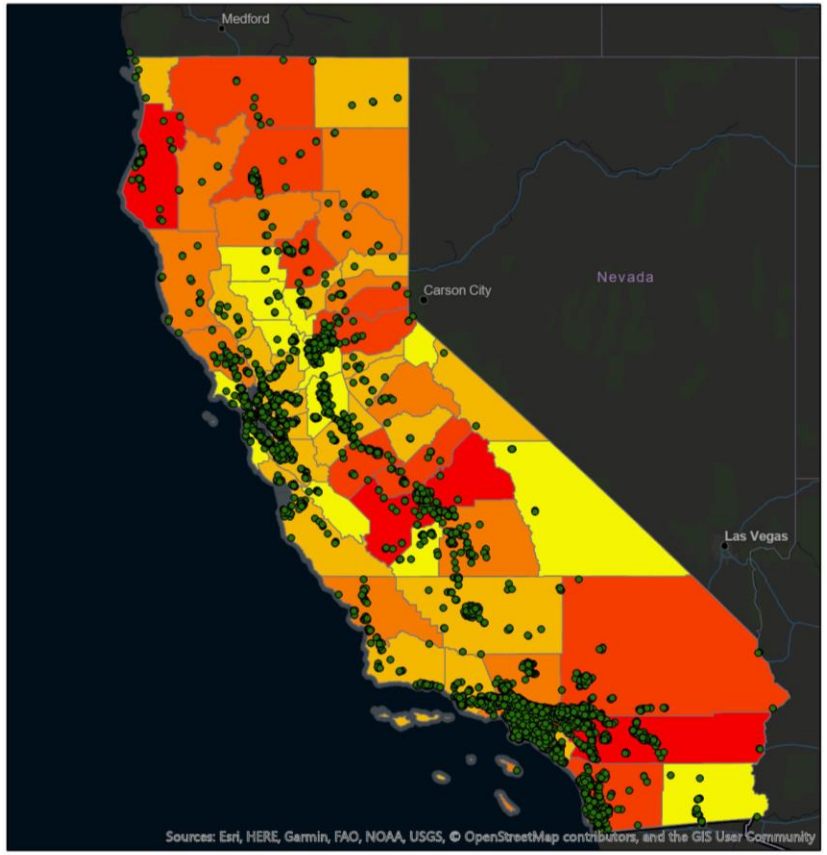
Fire Incident Count

- ≤41
- ≤114
- ≤185
- ≤320
- ≤614

2015 Fire Incidents by County and Fire Station Locations



2015 Fire Incidents by County and Healthcare Facility Locations



California Wildfires Dashboard

Code available on GitHub to run locally:

- https://github.com/gsharma14/DS-5500-California-Wildfires
- kepler.gl map URL
- Demo

2.

Modeling

Modeling - Methods

- Portugal
  - Fire Size (Regression)
- California
  - Data Enhancement
    - Climate data (NOAA)
    - Proximity Analysis
  - Fire Size (Regression)
  - Fire Cause (Classification)

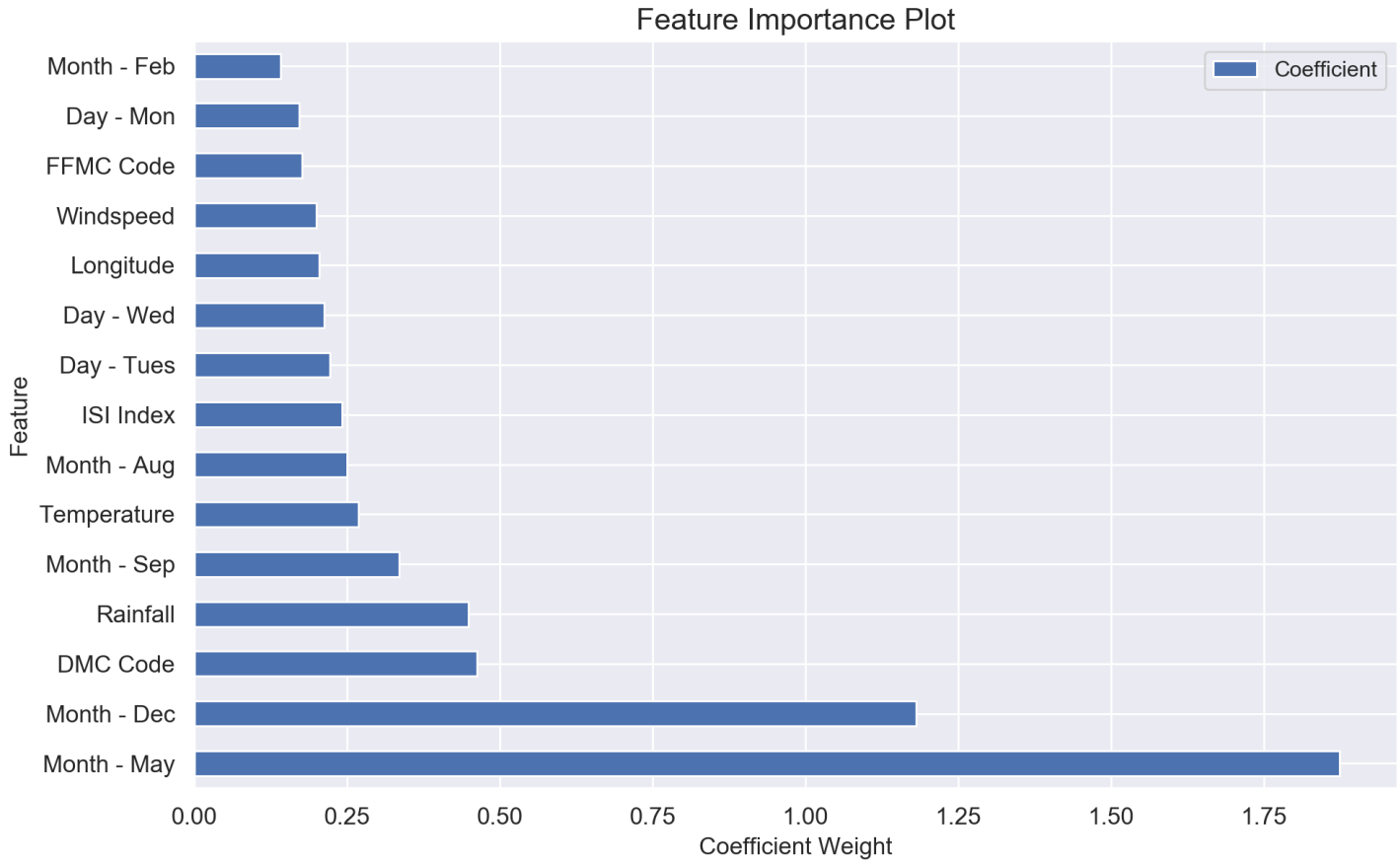
Portugal Dataset

Feature	Description	Type
Latitude/Longitude	Fire Location - Degrees	Continuous
Month	12 values	Nominal
Day of Week	7 values	Nominal
FFMC (Fine Fuel Moisture Code)	Fuel flammability/ease of ignition	Continuous
DMC (Duff Moisture Code)	Organic matter moisture content	Continuous
DC (Drought Code)	Indicator of seasonal drought	Continuous
ISI (Initial Spread Index)	Fire spread immediatley after ignition	Continuous
Temperature	Degrees – Celsius	Continuous
Relative Humidity (RH)	Percentage	Continuous
Wind Speed	km/h	Continuous
Rainfall	mm/m^2	Continuous
Area	Forest area burned (hectare)	Continuous

N = 517, Fire Weather Index (FWI) Measures

Fire Size Prediction – Portugal

Regression Model	Cross-validation MAE
Linear Regression	18.16
Random Forest	16.61
SVM (linear)	11.07
XGBoost	21.55
Dense Neural Network	24.36



California Dataset

Feature	Description	Type
Latitude/Longitude	Fire Location - Degrees	Continuous
Fire Year	2000-2015	Continuous
County	59 counties in California	Nominal
Fire Cause (STAT_CAUSE_DESCR)	13 values	Nominal
Fire Size	Acres	Continuous
Average Wind Speed (AWND)	Yearly county average, mph	Continuous
Average Temperature (TAVG)	Yearly county average, Fahrenheit	Continuous
Average Precipitation (PRCP)	Yearly county average, cm	Continuous
Nearest Fire Station	Meters to nearest fire station	Continuous
Above 90° F (DX90)	Days with > 90° F in county	Continuous
Burn Time	Difference between discovery date and containment date	Continuous

N = 53128, enhanced with NOAA data

Fire Size Prediction – California

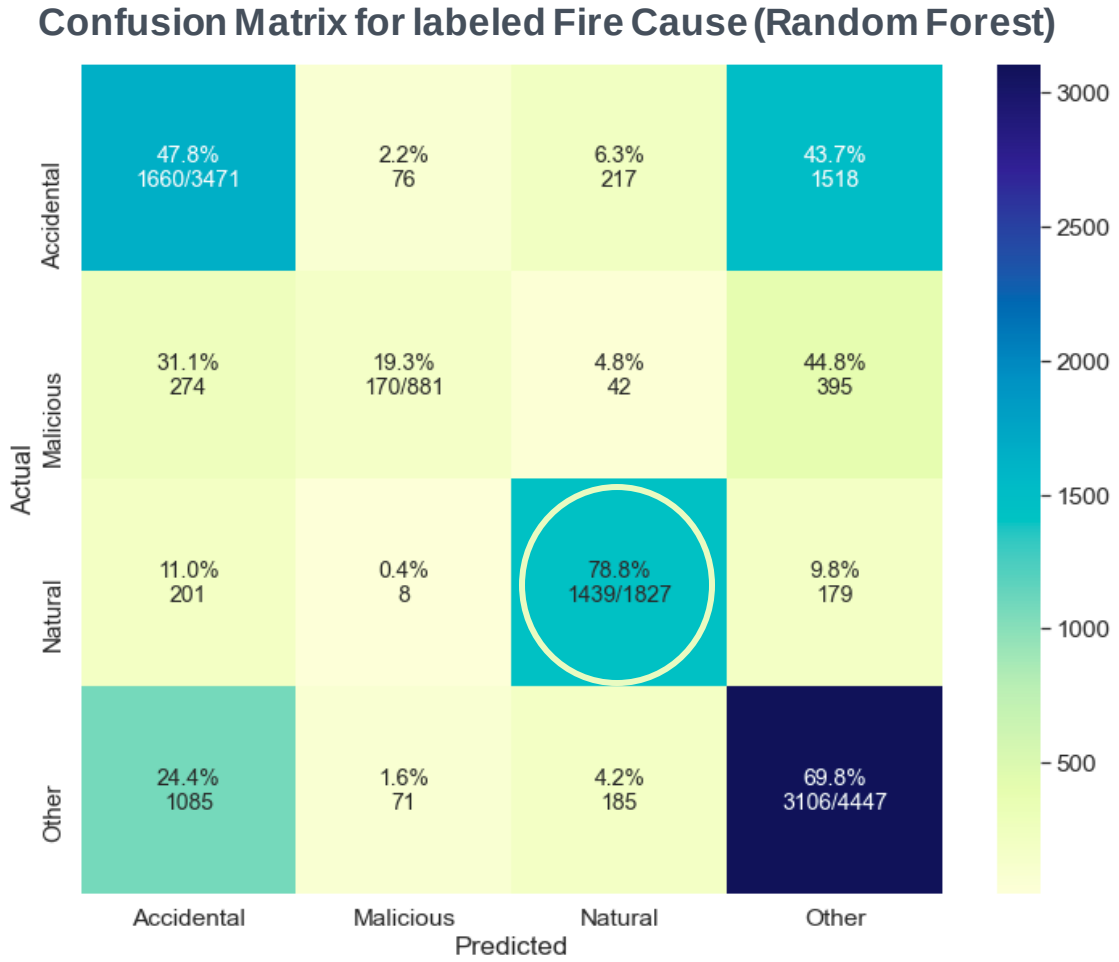
Regression Model	Test MAE
Linear Regression	230.9
Random Forest	161.4
XGBoost	183.9
SVM (linear)	120.4

Fire Cause Prediction – California (Multi-Class)

- Without Label
  - Random Forest Classification
- With Label
  - Decision Tree Classifier with Adaboost
  - Gaussian and Multinomial Naive Bayes
  - KNeighbours Classifier
  - Gradient Boosting Classifier
  - **Random Forest Classification *(best model score)***

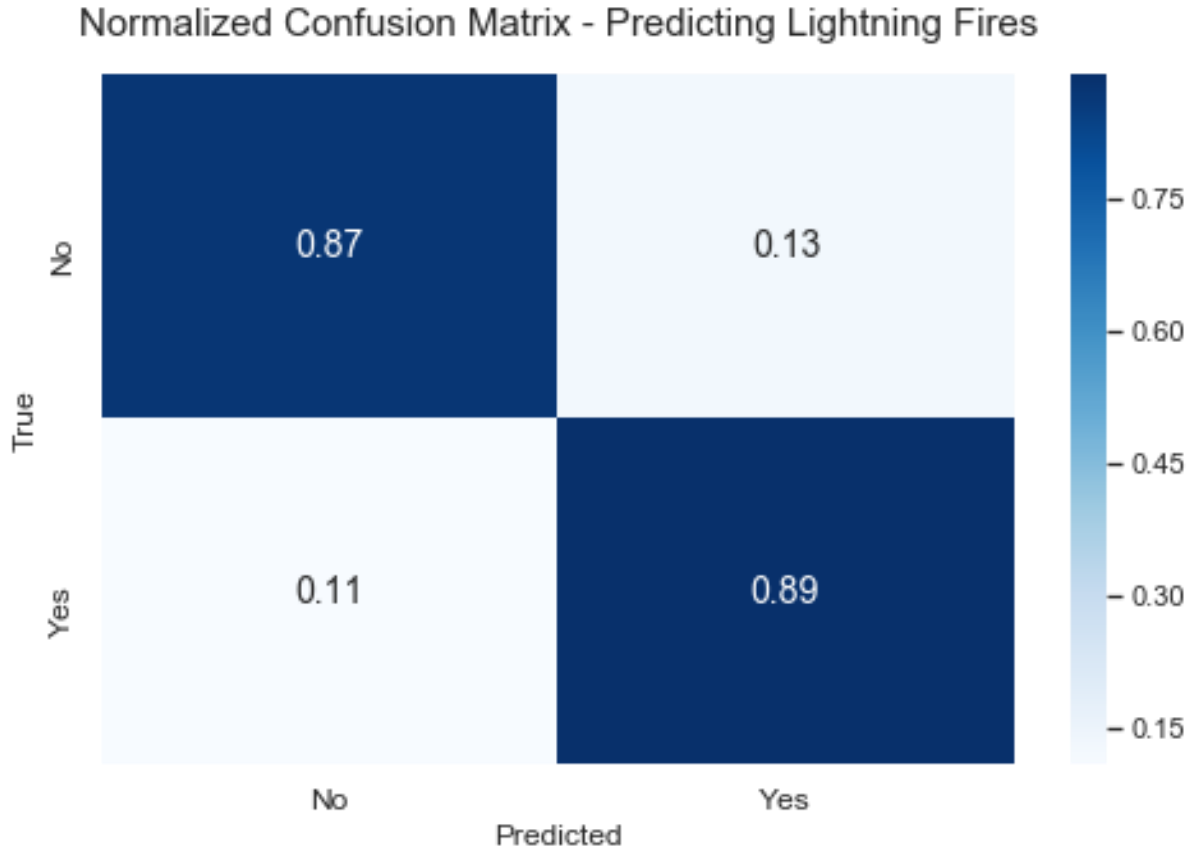
Fire Cause – Multi-class, with labels

CATEGORIES	LABEL
Natural: {Lightning}	Natural
Accidental: {Structure, Firework, Powerline, Railroad, Smoking, Children, Campfire, Equipment Use, Debris Burning}	Accidental
Malicious: {Arson}	Malicious
Miscellaneous/Missing	Other

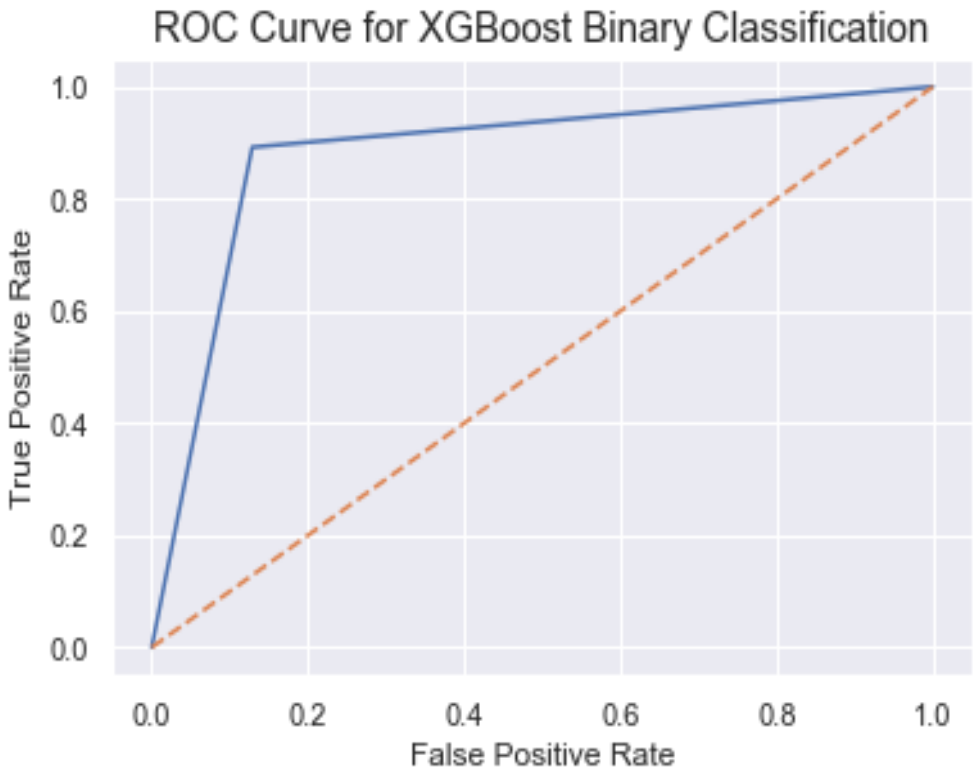


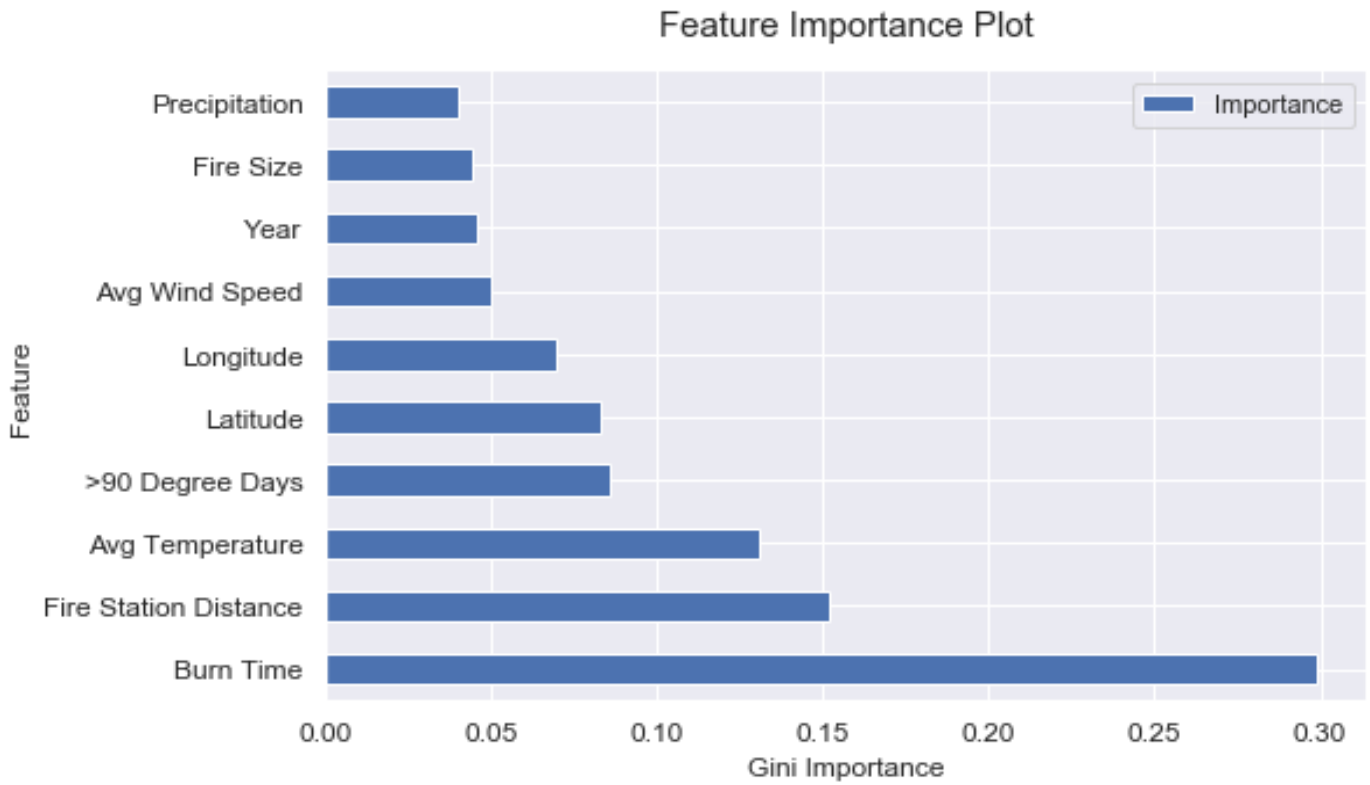
Fire Cause – Binary/Lightning

Classification Model	Test Accuracy	Test Precision	Test Recall
Logistic Regression	0.80	0.797	0.793
Random Forest	0.87	0.86	0.878
XGBoost	0.882	0.87	0.892



Fire Cause – Binary/Lightning





3.

Discussion

General Takeaways

- Project outcomes
  - Interactive dashboard
    - Spatial mapping
    - Fire cause/size trends
  - Modeling
    - Identified relevant features
- How can California better prepare for fires?
  - Improve access to resources and aid
  - Monitor weather conditions

TIMELINE

EXPLORING WILDFIRES

Future Work

- More recent data
- Granular climate data
- Multiple states
- Satellite imagery

Thanks!

Any questions?

Datasets

1. US Forest Service. (2020, April 29). National Interagency Fire Occurrence 1992-2015 (Feature Layer). https://enterprisecontentnew-usfs.hub.arcgis.com/datasets/e4d020cb51304d5194860d4464da7ba7_0/data?geometry=61.662%2C-2.200%2C54.279%2C76.163
2. UCI. (n.d.). UCI Machine Learning Repository: Forest Fires Data Set. UCI Machine Learning Repository. Retrieved 2020, from https://archive.ics.uci.edu/ml/datasets/Forest+Fires
3. National Oceanic and Atmospheric Administration (NOAA). (n.d.). Search | Climate Data Online (CDO) | National Climatic Data Center (NCDC). Climate Data Online (CDO). Retrieved October 2020, from https://www.ncdc.noaa.gov/cdo-web/search

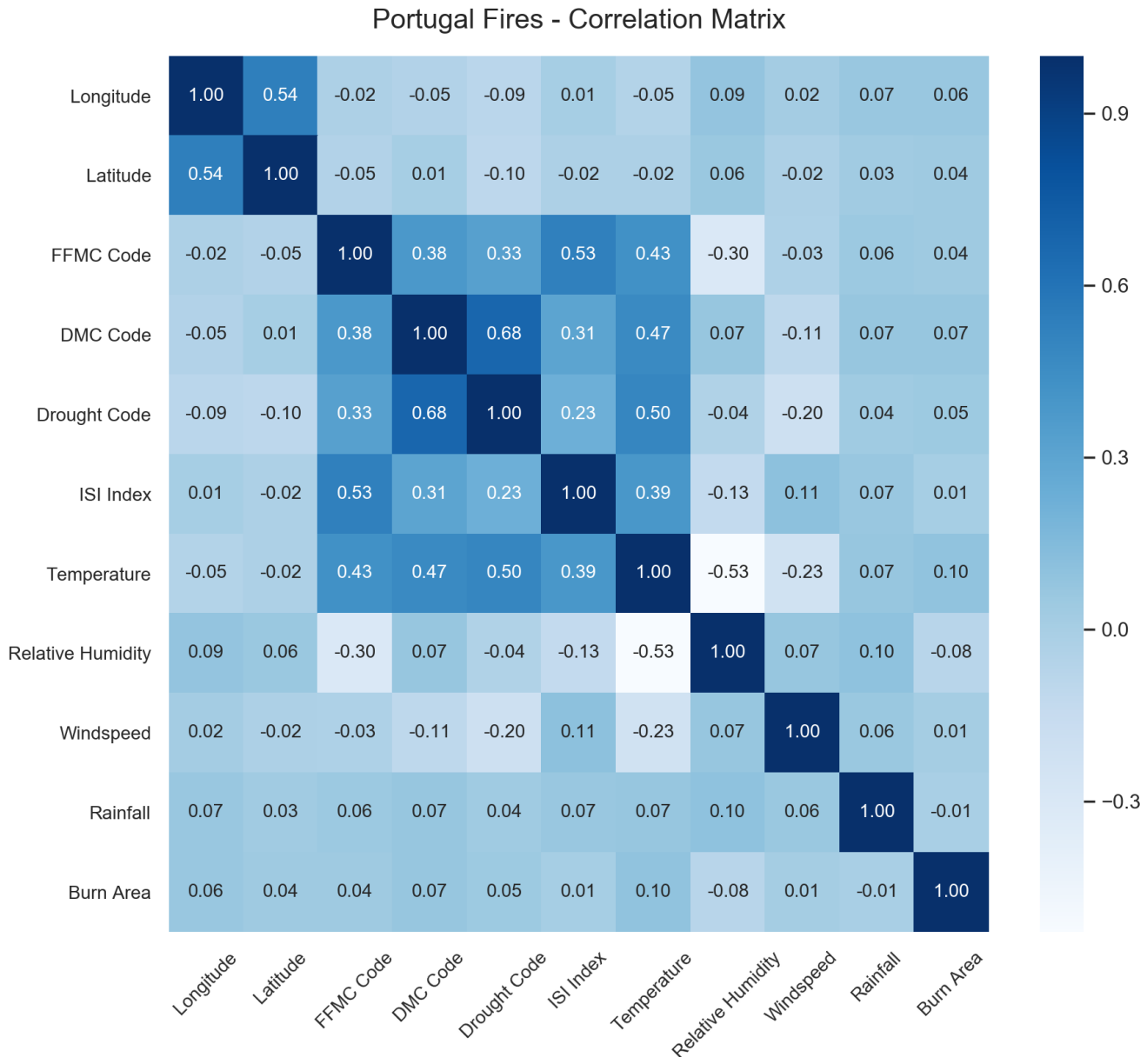
Additional Spatial Datasets

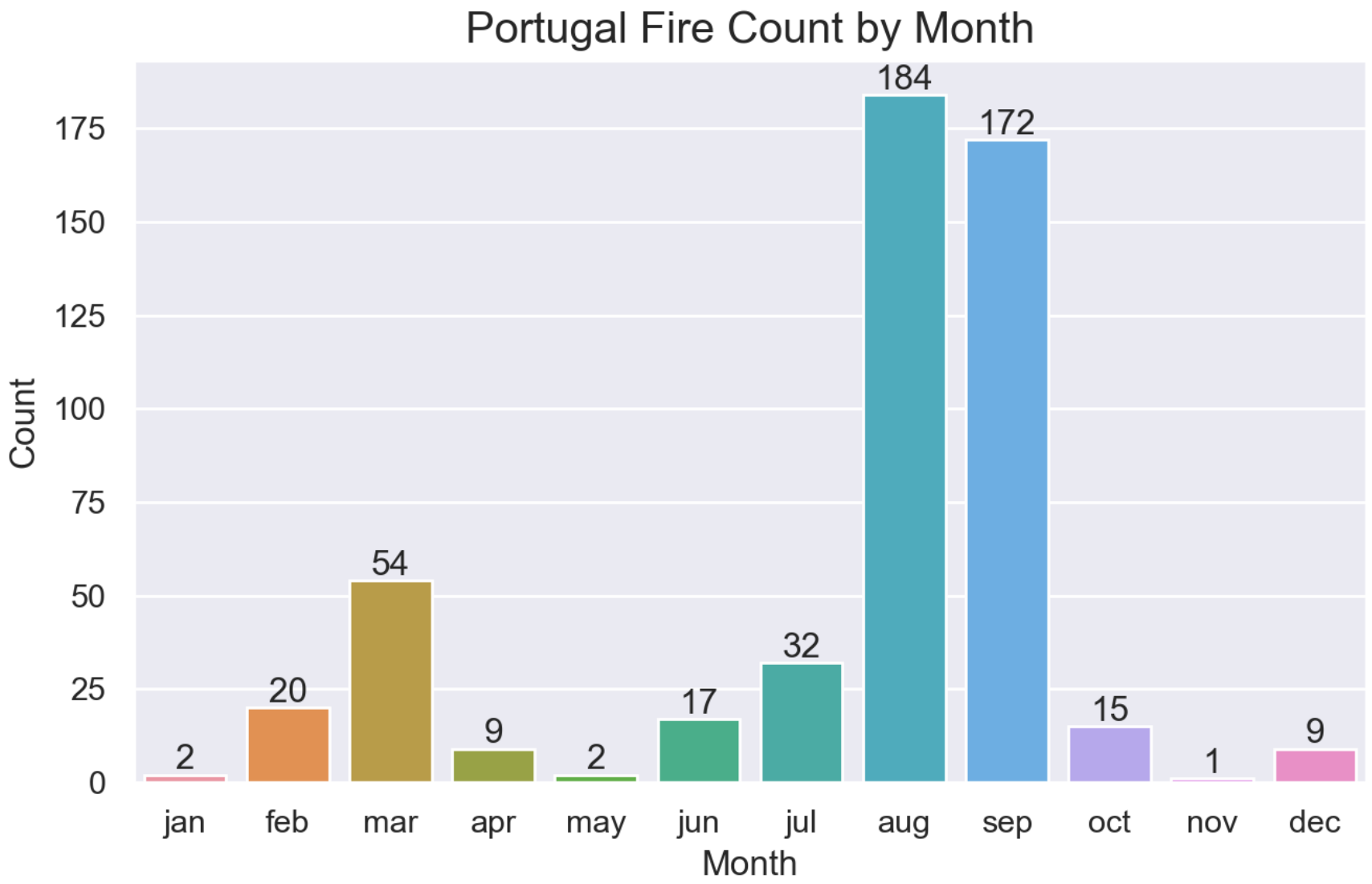
1. CA.gov. (2019, October 23). CA Geographic Boundaries - California Open Data. California Open Data Portal. https://data.ca.gov/dataset/ca-geographic-boundaries
2. ESRI. (n.d.). Living Atlas of the World | ArcGIS. ArcGIS. Retrieved October 2020, from https://livingatlas.arcgis.com/en/home/
3. CA Open Data - Department of Public Health. (2020, October 18). CDPH Licensing and Certification Healthcare Facilities. California State Geoportal. https://gis.data.ca.gov/datasets/CDPHDATA::cdph-licensing-and-certification-healthcare-facilities
4. Homeland Infrastructure Foundation-Level Data (HIFLD). (2020, September 11). Fire Stations. HIFLD Open Data. https://hifld-geoplatform.opendata.arcgis.com/datasets/fire-stations?geometry=-130.326%2C33.769%2C-104.442%2C39.917
5. US Census Bureau. (2019, August 15). TIGER/Line Shapefile, 2018, state, California, Primary and Secondary Roads State-based Shapefile - Data.gov. Data.Gov. https://catalog.data.gov/dataset/tiger-line-shapefile-2018-state-california-primary-and-secondary-roads-state-based-shapefile
6. CA.gov. (2020, August 11). California Electric Substation. California State Geoportal. https://gis.data.ca.gov/datasets/CAEnergy::california-electric-substation

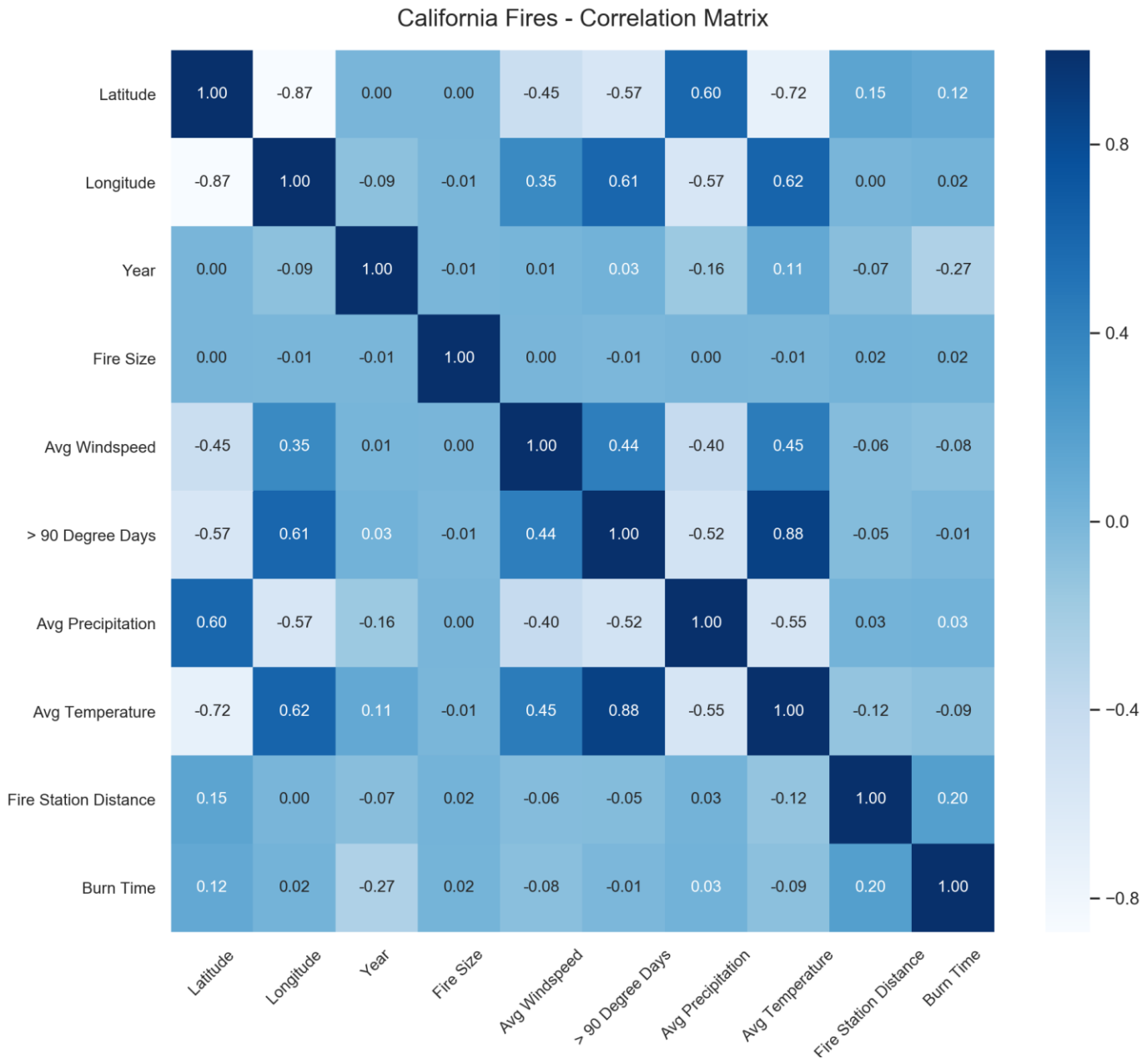
References

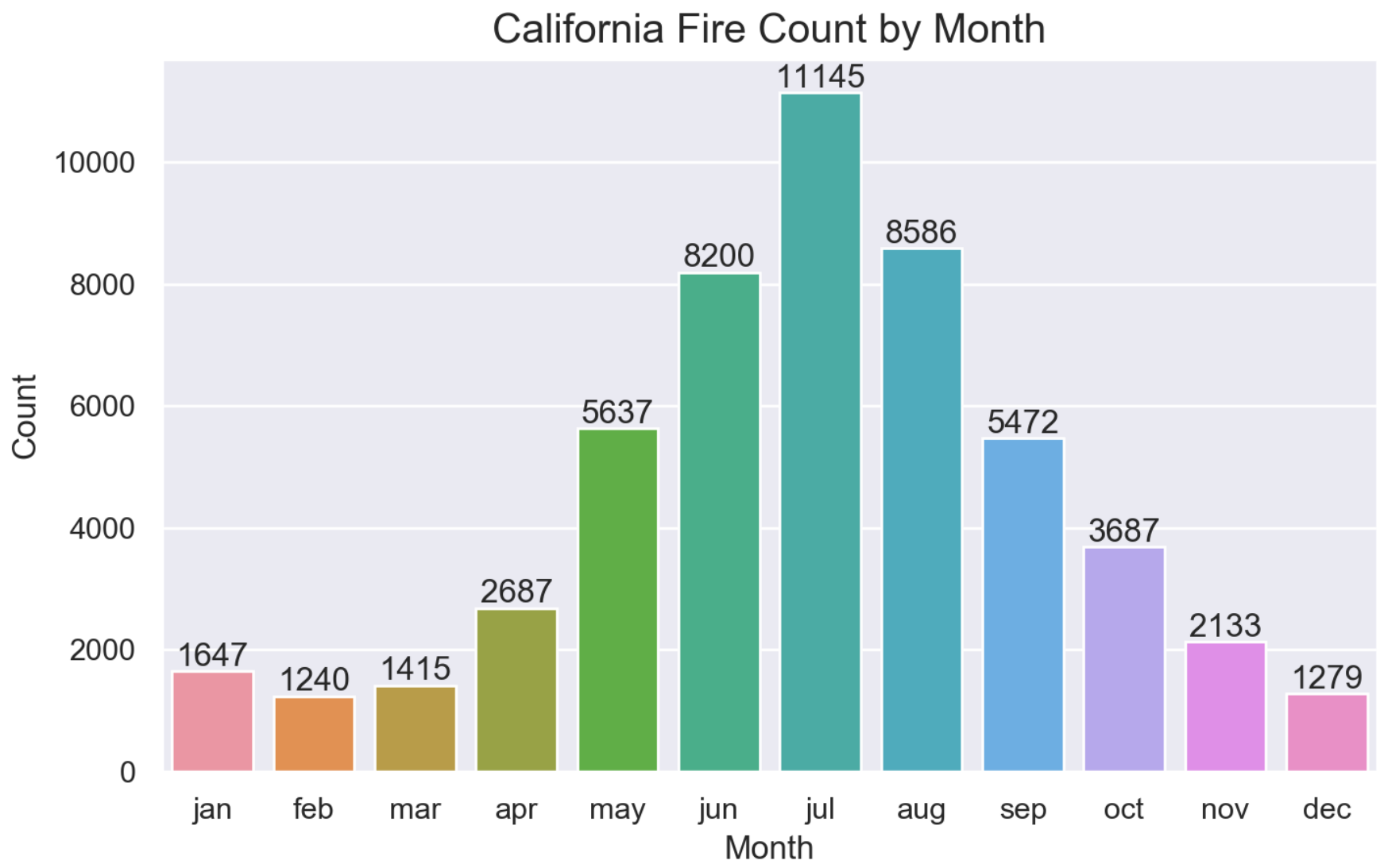
1. Center for Disaster Philanthropy (CPP). (2020, October 23). 2020 North American Wildfire Season. Center for Disaster Philanthropy. https://disasterphilanthropy.org/disaster/2020-california-wildfires/
2. Wikimedia. (n.d.). *2020 California Wildfires* [Map]. Wikimedia. https://upload.wikimedia.org/wikipedia/commons/8/85/2020_California_wildfires.png
3. Law, M., & Collins, A. (2019). Getting to Know ArcGIS Pro (Second ed.). Esri Press.
4. Bambrick, G. (2016, January 22). What is Hotspot Analysis? Geospatiality. https://glenbambrick.com/2016/01/21/what-is-hotspot-analysis/
5. kepler.gl. (2020). kepler.gl for Jupyter User Guide. https://docs.kepler.gl/docs/keplergl-jupyter
6. plotly. (n.d.). Part 2. Layout | Dash for Python Documentation | Plotly. Plotly | Dash. Retrieved 2020, from https://dash.plotly.com/layout
7. Stanford-Moore, A., & Moore, B. (2019). Wildfire Burn Area Prediction. Stanford University - CS229. http://cs229.stanford.edu/proj2019aut/data/assignment_308832_raw/26582553.pdf
8. ArcGIS. (n.d.). Historical Wildfire Analysis | ArcGIS for Developers. ArcGIS API for Python. Retrieved 2020, from https://developers.arcgis.com/python/sample-notebooks/historical-wildfire-analysis/
9. O'Neill, P. (2017, December 13). Predict the causes of wildfires using Python. Kaggle. https://www.kaggle.com/edhirif/predict-the-causes-of-wildfires-using-python

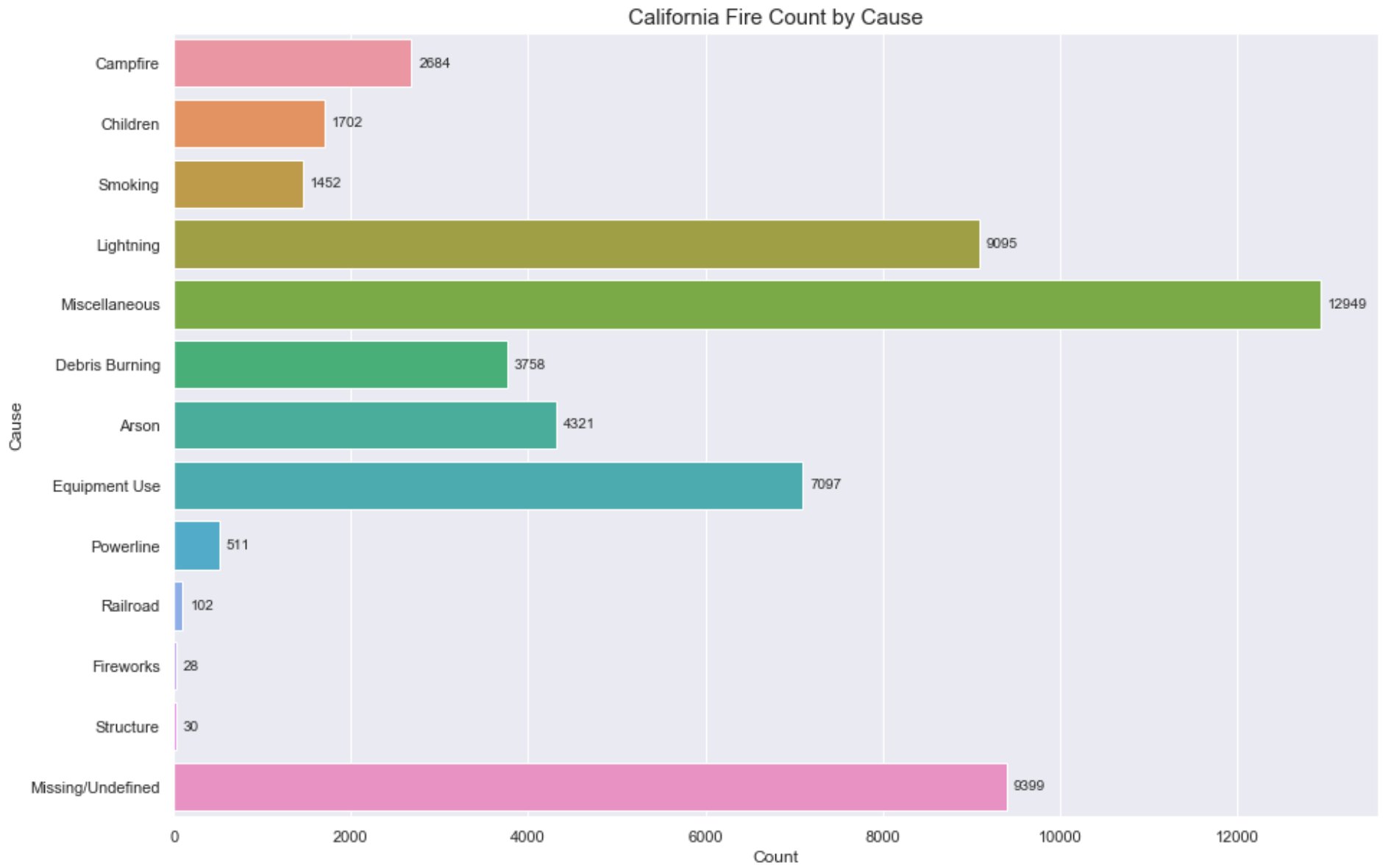
Appendix

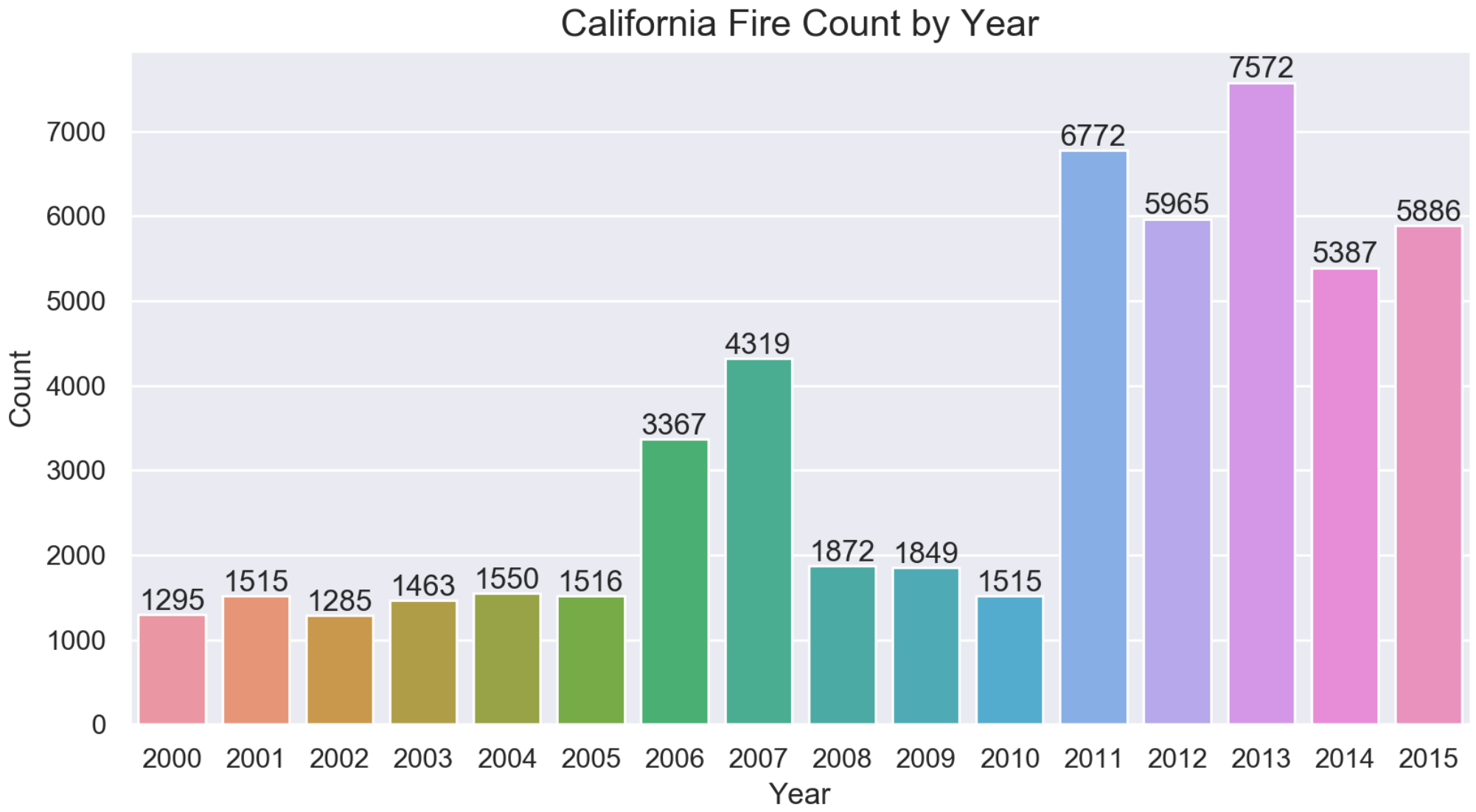

Portugal Fires - Correlation Matrix
Longitude 1.00 0.54 -0.02 -0.05 -0.09 0.01 -0.05 0.09 0.02 0.07 0.06
Latitude 0.54 1.00 -0.05 0.01 -0.10 -0.02 -0.02 0.06 -0.02 0.03 0.04
FFMC Code -0.02 -0.05 1.00 0.38 0.33 0.53 0.43 -0.30 -0.03 0.06 0.04
DMC Code -0.05 0.01 0.38 1.00 0.68 0.31 0.47 0.07 -0.11 0.07 0.07
Drought Code -0.09 -0.10 0.33 0.68 1.00 0.23 0.50 -0.04 -0.20 0.04 0.05
ISI Index 0.01 -0.02 0.53 0.31 0.23 1.00 0.39 -0.13 0.11 0.07 0.01
Temperature -0.05 -0.02 0.43 0.47 0.50 0.39 1.00 -0.53 -0.23 0.07 0.10
Relative Humidity 0.09 0.06 -0.30 0.07 -0.04 -0.13 -0.53 1.00 0.07 0.10 -0.08
Windspeed 0.02 -0.02 -0.03 -0.11 -0.20 0.11 -0.23 0.07 1.00 0.06 0.01
Rainfall 0.07 0.03 0.06 0.07 0.04 0.07 0.07 0.10 0.06 1.00 -0.01
Burn Area 0.06 0.04 0.04 0.07 0.05 0.01 0.10 -0.08 0.01 -0.01 1.00
Longitude
Latitude
FFMC Code
DMC Code
Drought Code
ISI Index
Temperature
Relative Humidity
Windspeed
Rainfall
Burn Area
0.9
0.6
0.3
0.0
−0.3

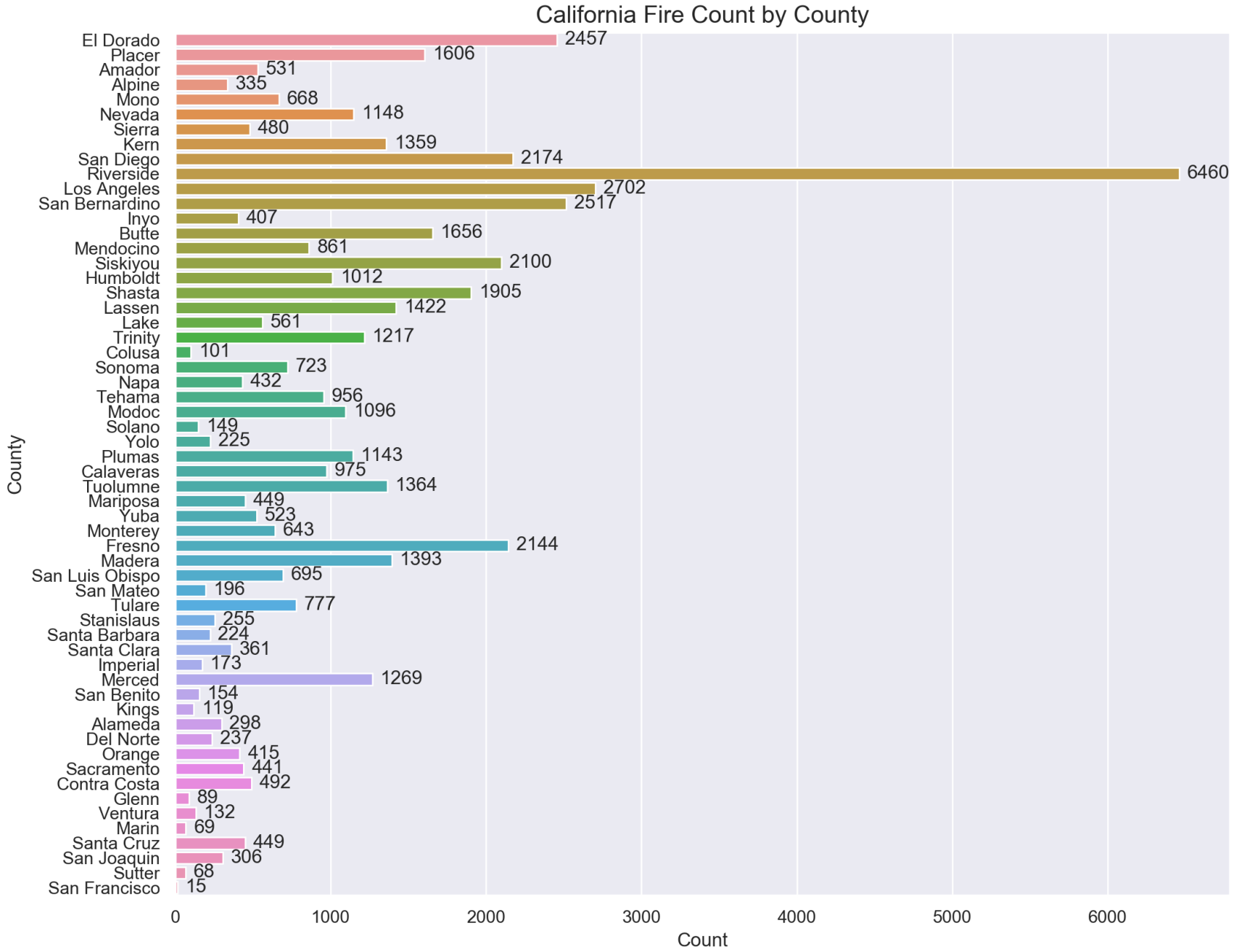

Portugal Fire Count by Month
Count
0
25
50
75
100
125
150
175
2
20
54
9
2
17
32
184
172
15
1
9
jan
feb
mar
apr
may
jun
jul
aug
sep
oct
nov
dec
Month


California Fires - Correlation Matrix
Latitude 1.00 -0.87 0.00 0.00 -0.45 -0.57 0.60 -0.72 0.15 0.12
Longitude -0.87 1.00 -0.09 -0.01 0.35 0.61 -0.57 0.62 0.00 0.02
Year 0.00 -0.09 1.00 -0.01 0.01 0.03 -0.16 0.11 -0.07 -0.27
Fire Size 0.00 -0.01 -0.01 1.00 0.00 -0.01 0.00 -0.01 0.02 0.02
Avg Windspeed -0.45 0.35 0.01 0.00 1.00 0.44 -0.40 0.45 -0.06 -0.08
> 90 Degree Days -0.57 0.61 0.03 -0.01 0.44 1.00 -0.52 0.88 -0.05 -0.01
Avg Precipitation 0.60 -0.57 -0.16 0.00 -0.40 -0.52 1.00 -0.55 0.03 0.03
Avg Temperature -0.72 0.62 0.11 -0.01 0.45 0.88 -0.55 1.00 -0.12 -0.09
Fire Station Distance 0.15 0.00 -0.07 0.02 -0.06 -0.05 0.03 -0.12 1.00 0.20
Burn Time 0.12 0.02 -0.27 0.02 -0.08 -0.01 0.03 -0.09 0.20 1.00
Latitude
Longitude
Year
Fire Size
Avg Windspeed
> 90 Degree Days
Avg Precipitation
Avg Temperature
Fire Station Distance
Burn Time
0.8
0.4
0.0
−0.4
−0.8


California Fire Count by Month
Count
0
2000
4000
6000
8000
10000
1647
1240
1415
2687
5637
8200
11145
8586
5472
3687
2133
1279
jan
feb
mar
apr
may
jun
jul
aug
sep
oct
nov
dec
Month


California Fire Count by Cause
Campfire
Children
Smoking
Lightning
Miscellaneous
Debris Burning
Arson
Equipment Use
Powerline
Railroad
Fireworks
Structure
Missing/Undefined
2684
1702
1452
9095
12949
3758
4321
7097
511
102
28
30
9399
Cause
0
2000
4000
6000
8000
10000
12000
Count


California Fire Count by Year
Count
Year
0
1000
2000
3000
4000
5000
6000
7000
1295
1515
1285
1463
1550
1516
3367
4319
1872
1849
1515
6772
5965
7572
5387
5886
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015


California Fire Count by County
El Dorado 2457
Placer 1606
Amador 531
Alpine 335
Mono 668
Nevada 1148
Sierra 480
Kern 1359
San Diego 2174
Riverside 6460
Los Angeles 2702
San Bernardino 2517
Inyo 407
Butte 1656
Mendocino 861
Siskiyou 2100
Humboldt 1012
Shasta 1905
Lassen 1422
Lake 561
Trinity 1217
Colusa 101
Sonoma 723
Napa 432
Tehama 956
Modoc 1096
Solano 149
Yolo 225
Plumas 1143
Calaveras 975
Tuolumne 1364
Mariposa 449
Yuba 523
Monterey 643
Fresno 2144
Madera 1393
San Luis Obispo 695
San Mateo 196
Tulare 777
Stanislaus 255
Santa Barbara 224
Santa Clara 361
Imperial 173
Merced 1269
San Benito 154
Kings 119
Alameda 298
Del Norte 237
Orange 415
Sacramento 441
Contra Costa 492
Glenn 89
Ventura 132
Marin 69
Santa Cruz 449
San Joaquin 306
Sutter 68
San Francisco 15
County
0
1000
2000
3000
4000
5000
6000
Count

Fire Cause – Multi-class, without labels