FIRE RISK

Marta Jesse CM Kira

Team Burnout

FIRE IS A CLEAR AND PRESENT DANGER

2020 was the most active fire season in the Western United States's recorded history.

Worst Fire Season in CA, Worst in a decade AZ, Most destructive OR, Largest Fires on Record in WA and CO

2020 Fire Damage Report

Land Up in Flame

10.2 million acres of land burned.

Lives Lost

46 (32 in California, 11 in Oregon, 1 in Washington, 2 in Colorado)

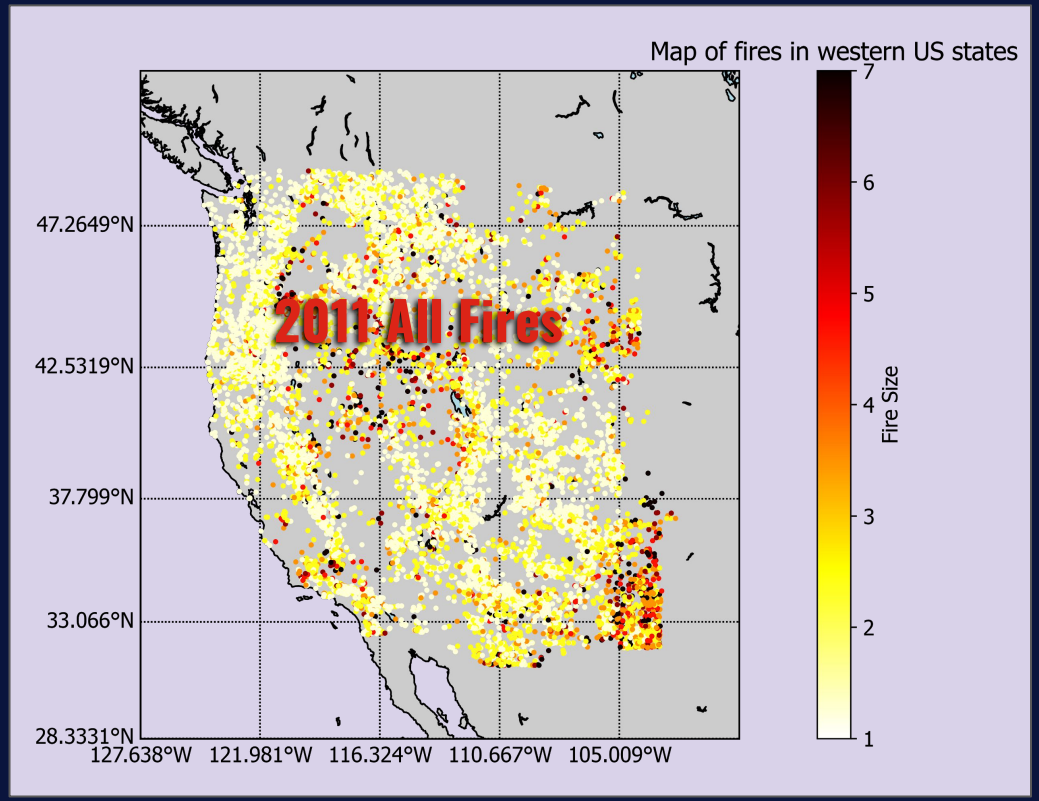
Carbon Emissions

Oregon and California broke records for carbon emissions

Properties Lost

10,488 structures in CA
$19.884 billion Cost

COVERAGE AREA



- Arizona
- California
- Colorado
- Idaho
- Oregon
- New Mexico
- Nevada
- Montana
- Utah
- Washington
- Wyoming

OUR DATASETS

WEATHER

120 years' worth of precipitation, temperatures, and drought indices

FIRES

24 years of fires' history, incl. cause, location, final size (in acres), and dates

CLEANING CHALLENGES

DATA CLEANING

Cleaning our data:

- Standard stuff like lowercase columns etc.
- Very careful about looking at **NaNs**
- New Challenge: Date Time - **Julian to Gregorian**
- Limited States to the **West**
- Limited years to **1987 - 2015**
- Removed Outliers - Fires that burned **over 5,000 days**

COMBINING DATA SETS

Combining our data:

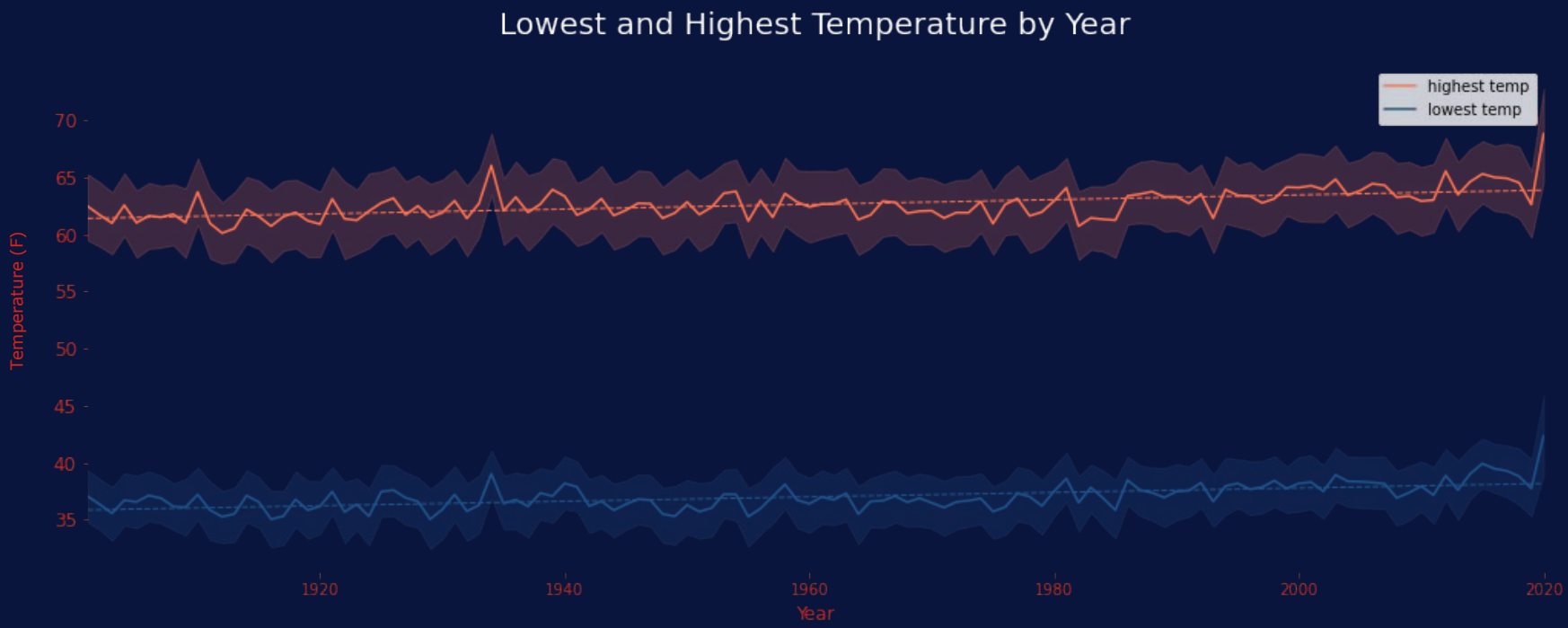
- Data Size was a limitation
- Set up **AWS S3** bucket
- Set up **AWS CLI** for the team
- Merge file difficulties - **memory**
- Unable to get all data that we wanted
- SQL to pkl to csv fun

CHALLENGES

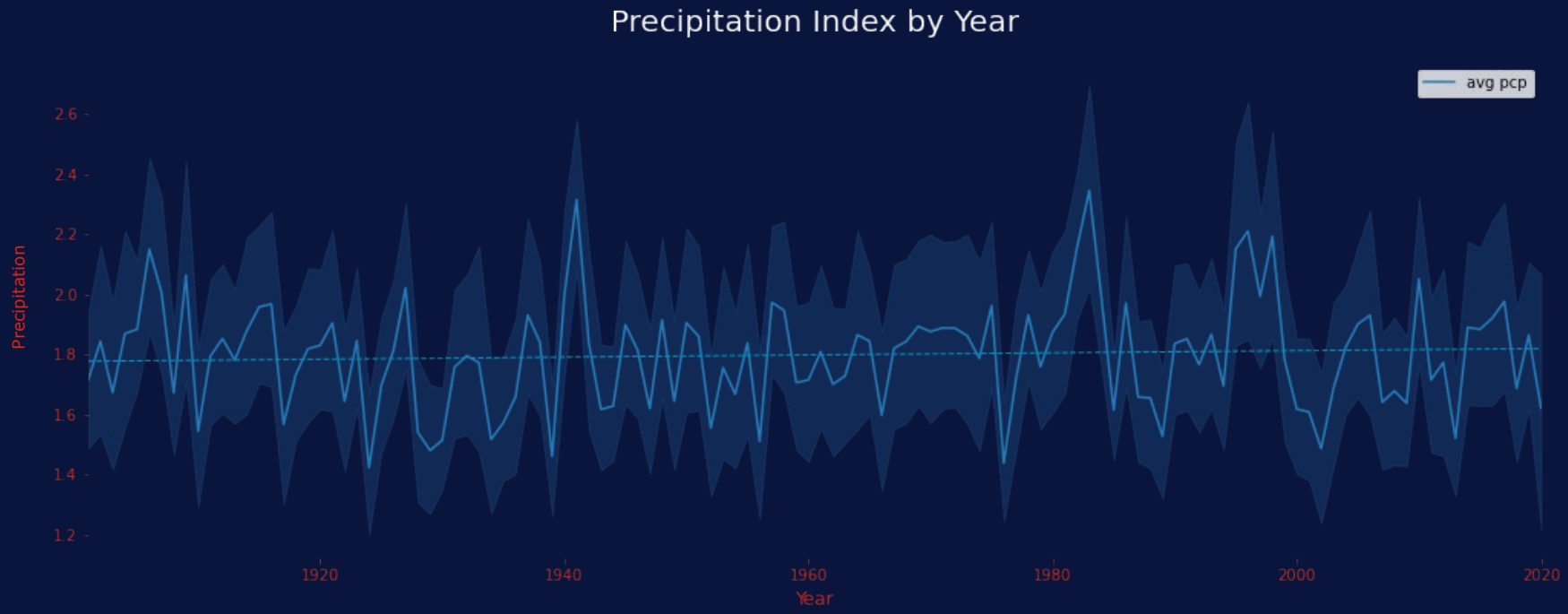
Difficulties with data:

- We realized that adding **foliage** and **wind** data would have been extremely helpful.
- Briefly considered trying to do something with **lightning**. - That consideration was quickly squashed.
- Limited data sets - Fire data only had **1992 - 2015**
- **Resource** limitations (processing power!)

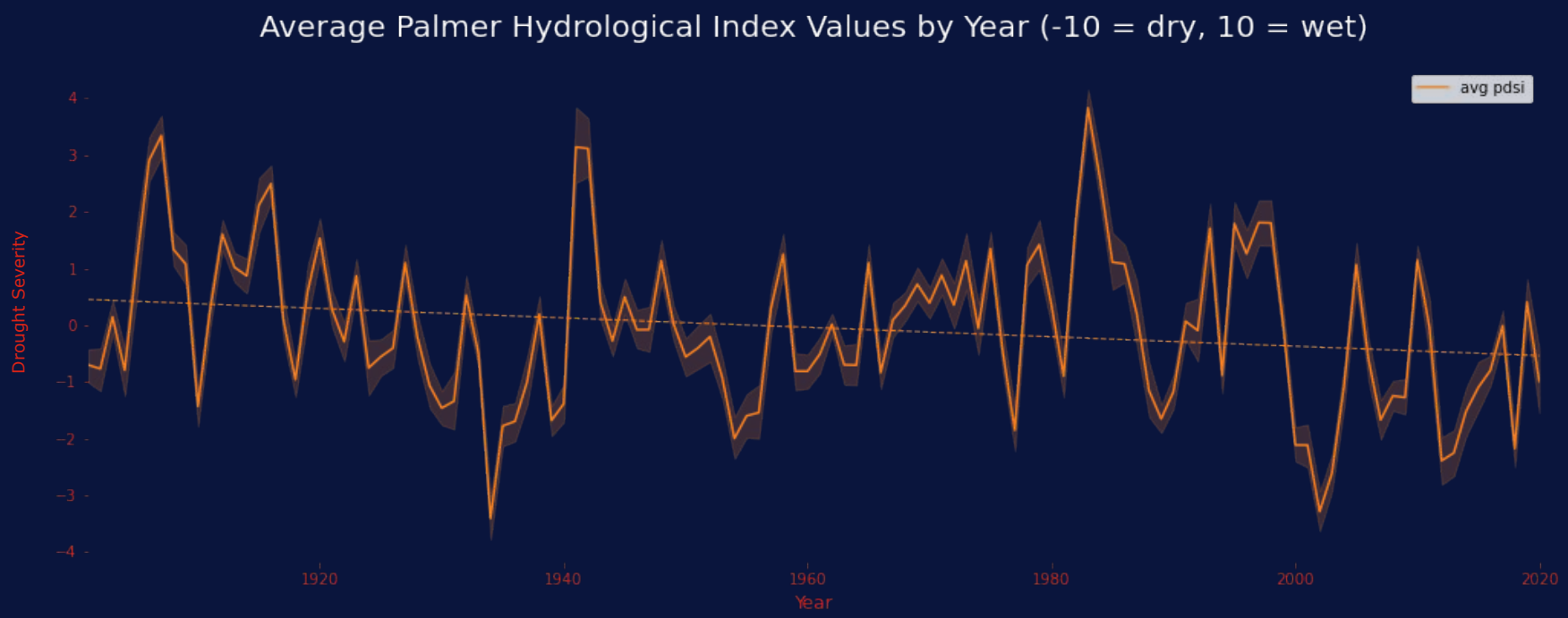
WEATHER TRENDS ARE SLOW BUT POWERFUL



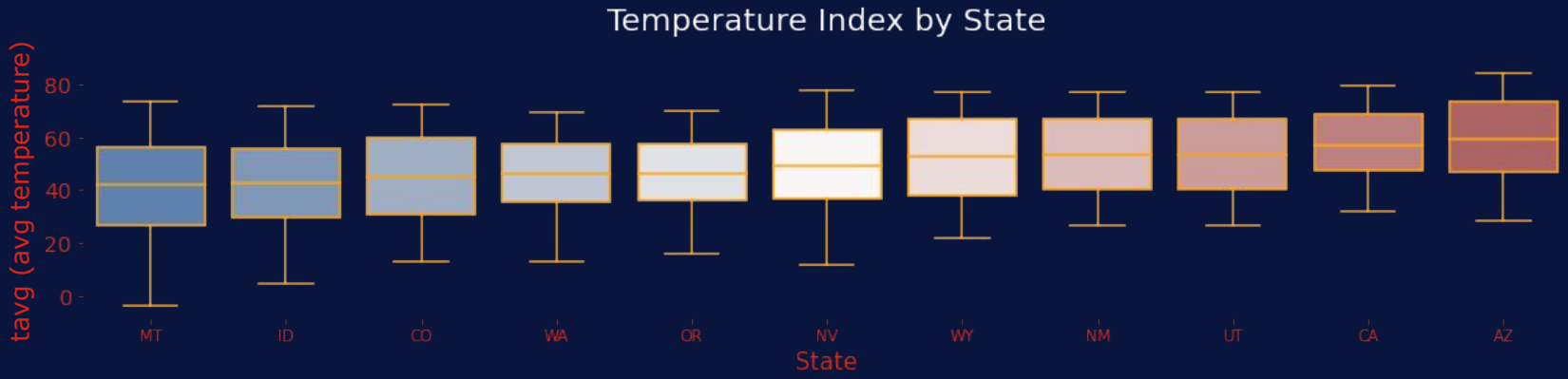
WEATHER TRENDS ARE SLOW BUT POWERFUL

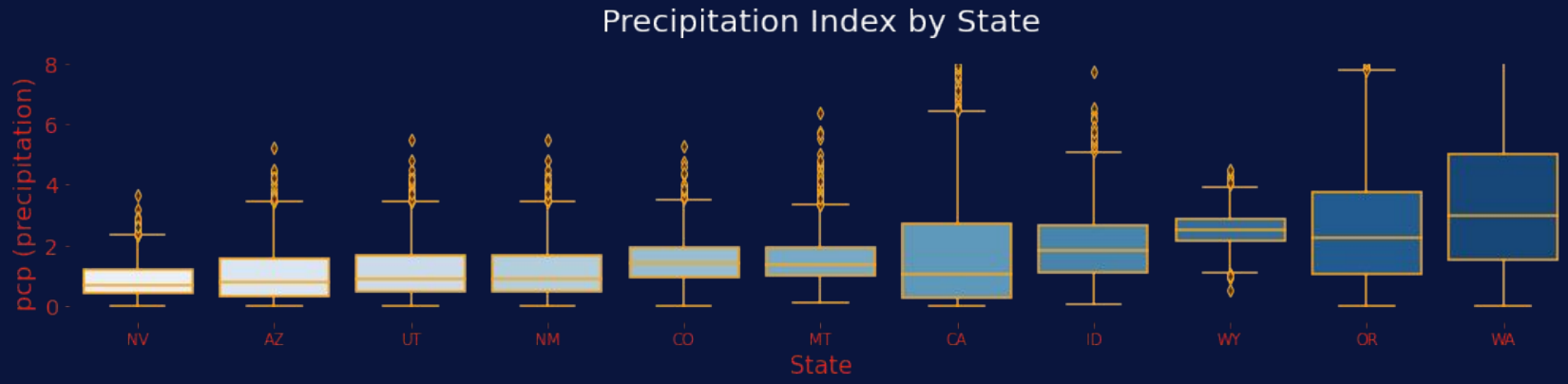


WEATHER TRENDS ARE SLOW BUT POWERFUL

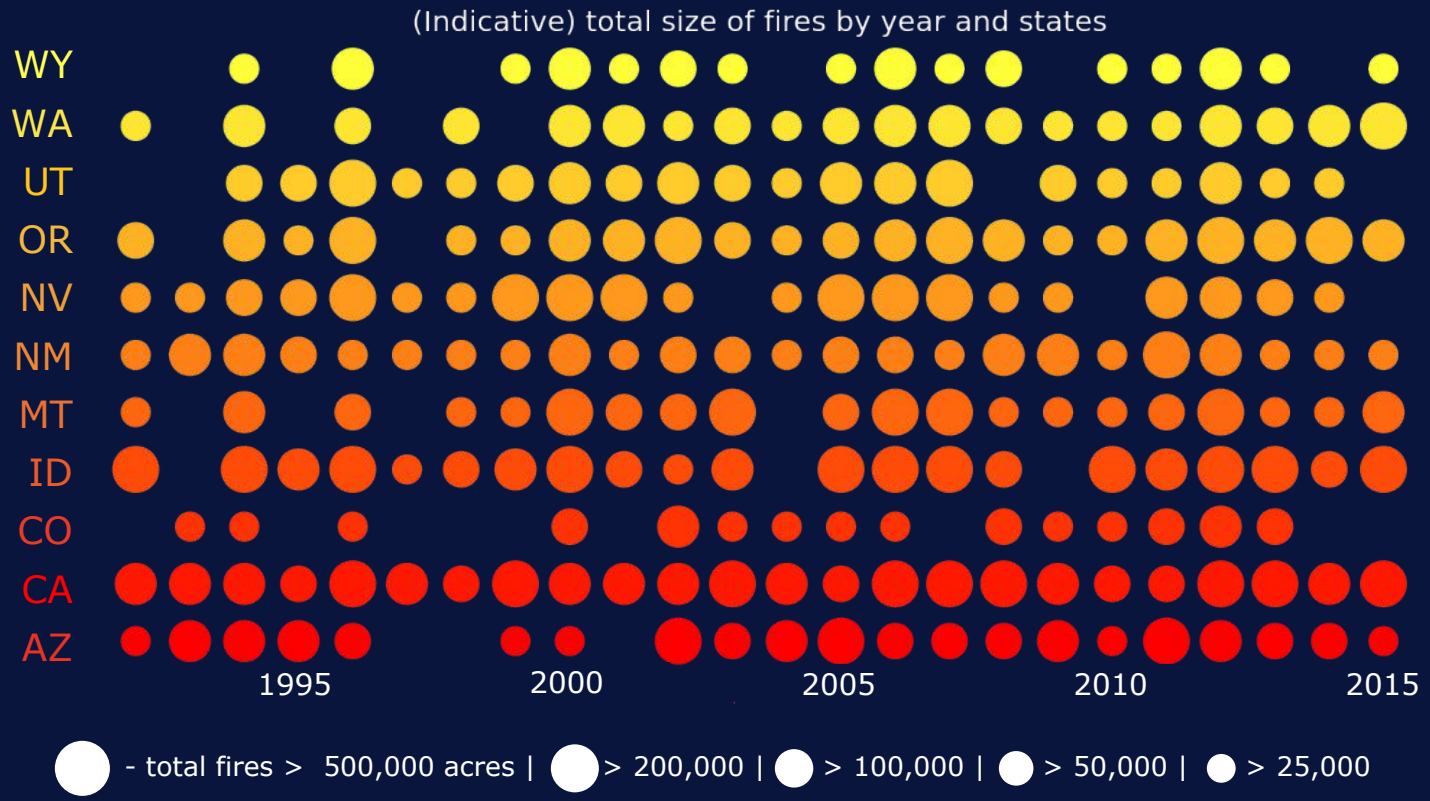


NOT ALL STATES ARE CREATED EQUAL

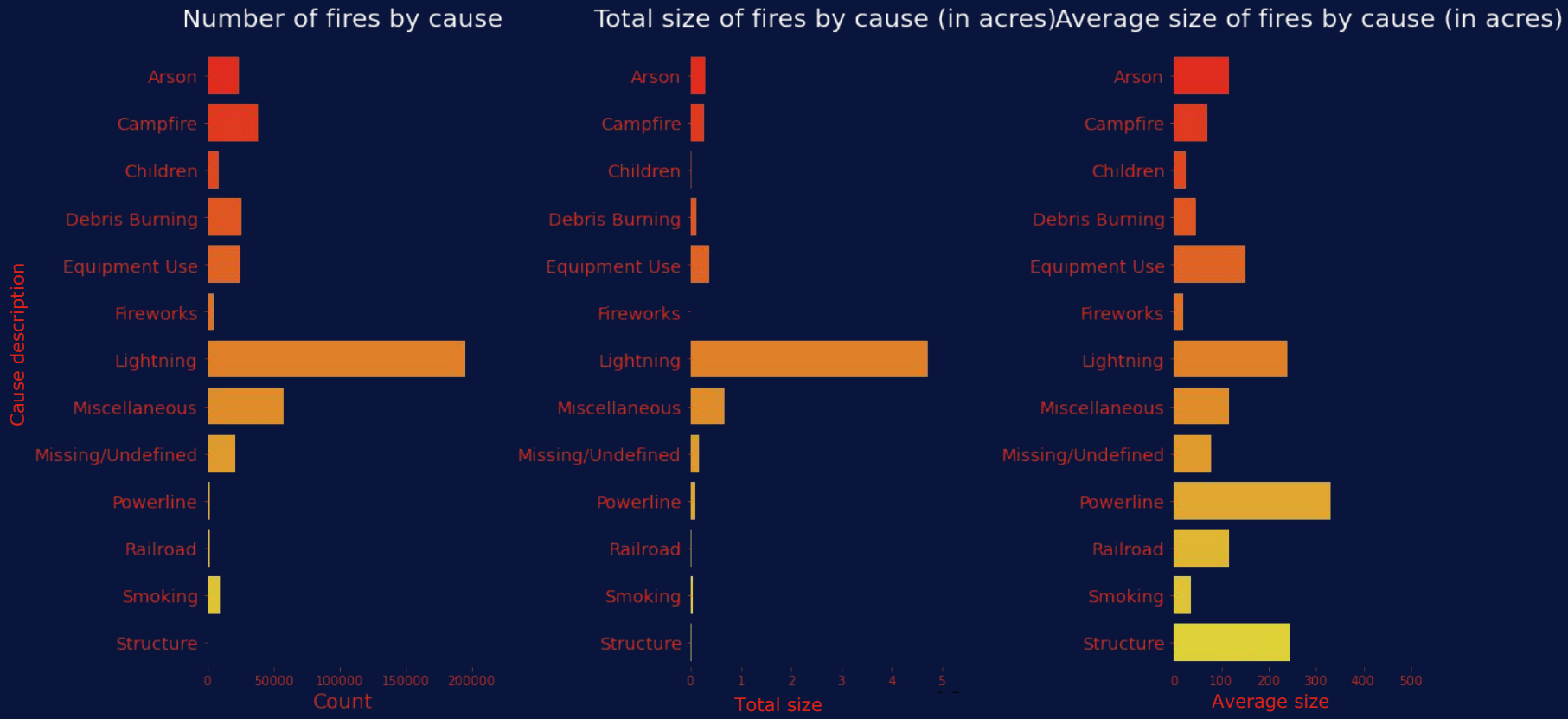




... AND NOT ALL STATES FEEL THE BURN EQUALLY

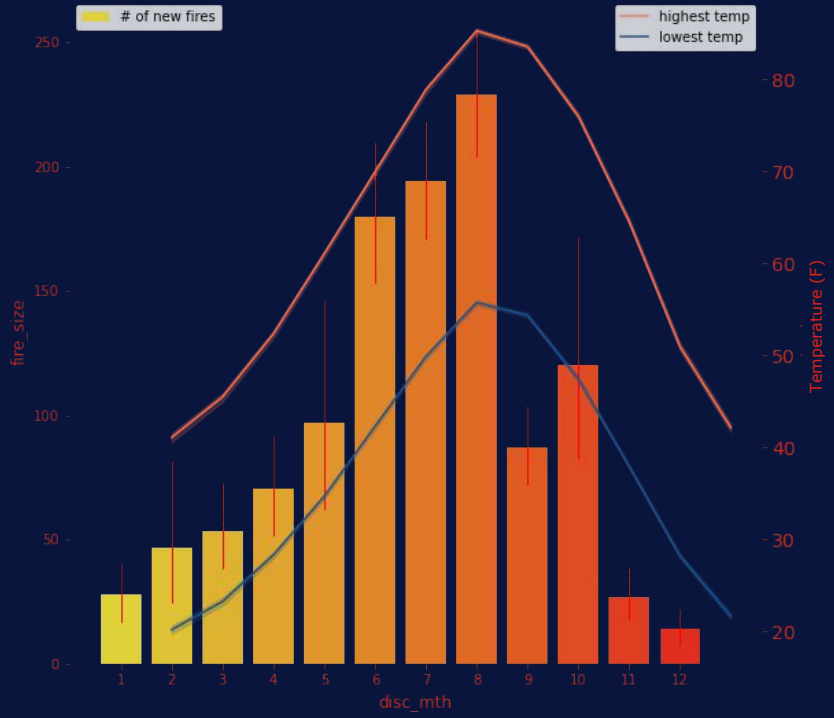


WE DIDN'T START THE FIRE... SO WHO DID?

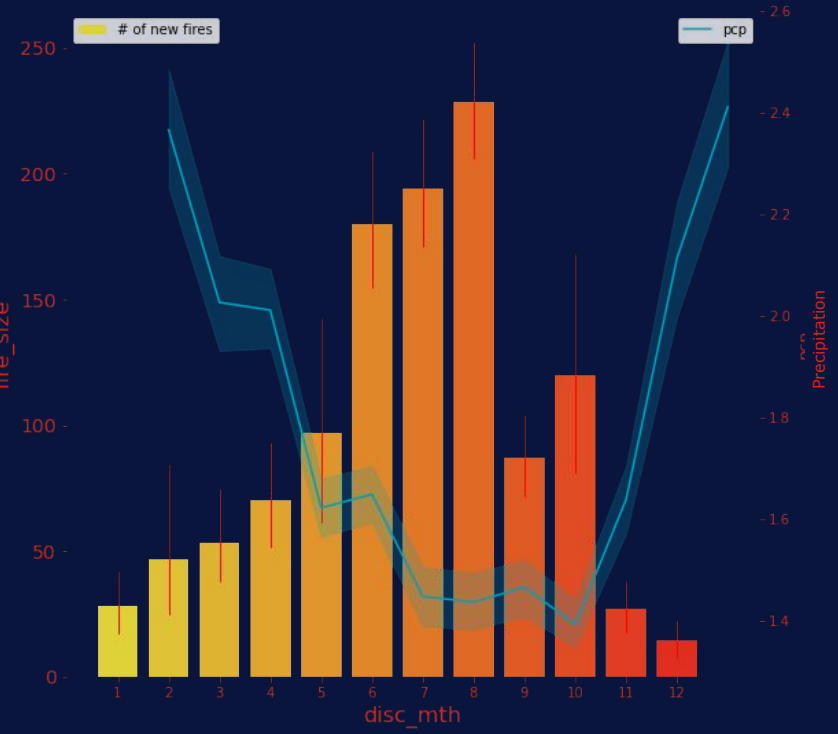


HIGH TEMPS + LOW PRECIPITATION = FIRE

Avg temperature ranges vs avg final size of fires by month of discovery



Avg precipitation index vs avg final size of fires by month of discovery



MODELING

REGRESSION

- Not the best fit model type for the scope of this project.
- Ridge Regression came out at around 15% R^2

THE MODEL IMPROVEMENT PROCESS

- Classification
- Bootstrapping
- Clustering
- Trailing Averages
- Class Weighting

CLASSIFICATION

RANDOM FOREST

- Purpose: feature importance

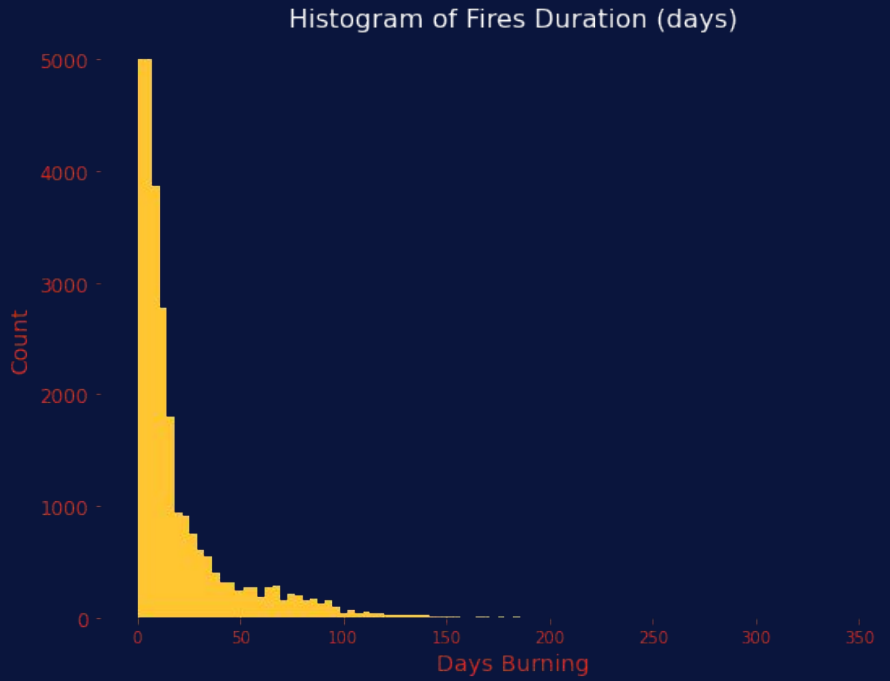
SVM

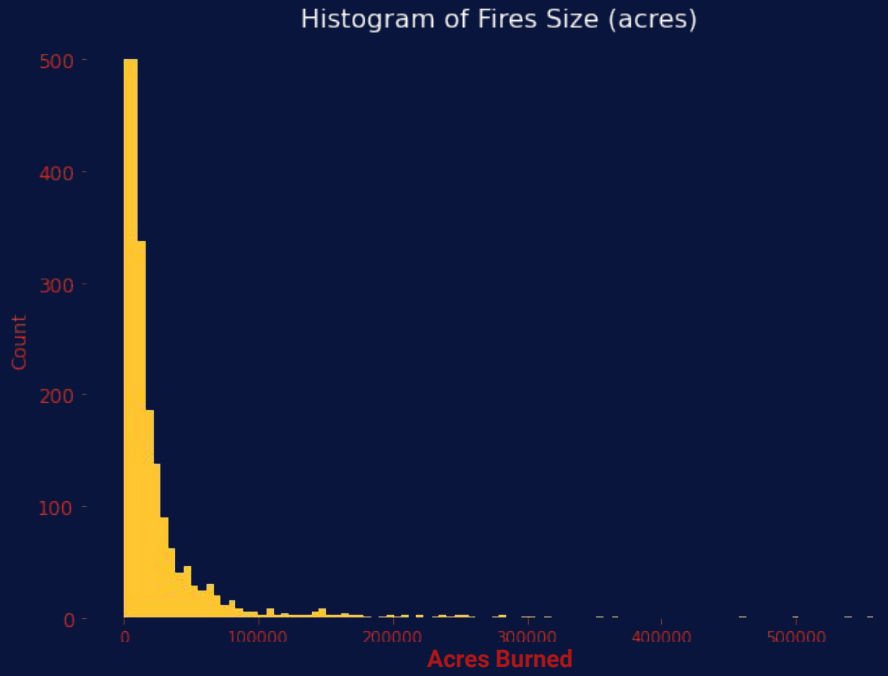
- Support Vector Machine with bootstrapped data

NEURAL NETWORK

- The Neural Network had the highest Recall Rate across the most important classes.
- This is the best predictive model.

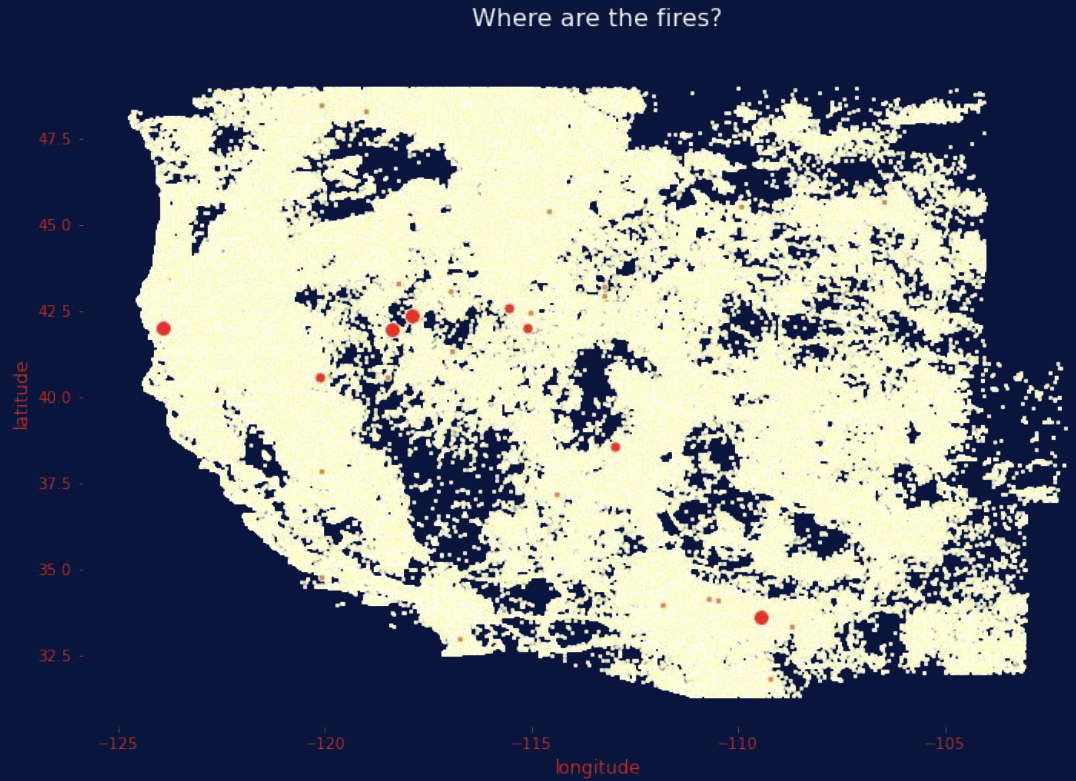
OUR CLASSES WERE MASSIVELY IMBALANCED



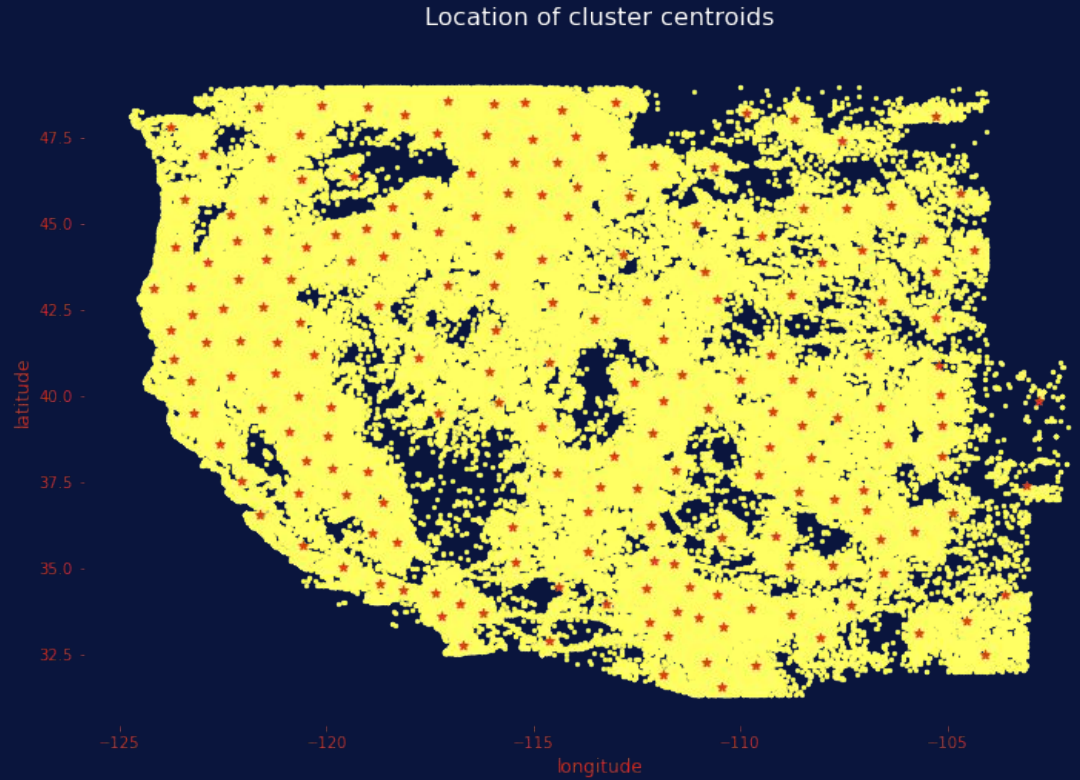


TARGET LABELS: "A" - fires smaller than 0.25 acres, "B" - 9.9 acres, "C" - 99.9 acres, "D" - 299 acres , "E" - 999 acres , "F" - 4999 acres, and "G" - 5000+ acres.

WHY WE CLUSTERED LATITUDE & LONGITUDE DATA



WHERE DID THE CLUSTER CENTROIDS LAND?



THE CHOSEN MODELS

NEURAL NETWORK CLASSIFIER

Purpose: predictive power

Optimized for Recall score, not Accuracy. Our highest Accuracy model was >70%. The chosen model came through at closer to 50% Accuracy, but a much higher recall rate.

Wildfire Class by Size and Recall:

	Modeling	Baseline
A	59%	62%
B	1%	29%
C	20%	6%
D	50%	2%
E	58%	1%
F	59%	2%
G	86%	.07%

RANDOM FOREST CLASSIFIER

Purpose: feature importance

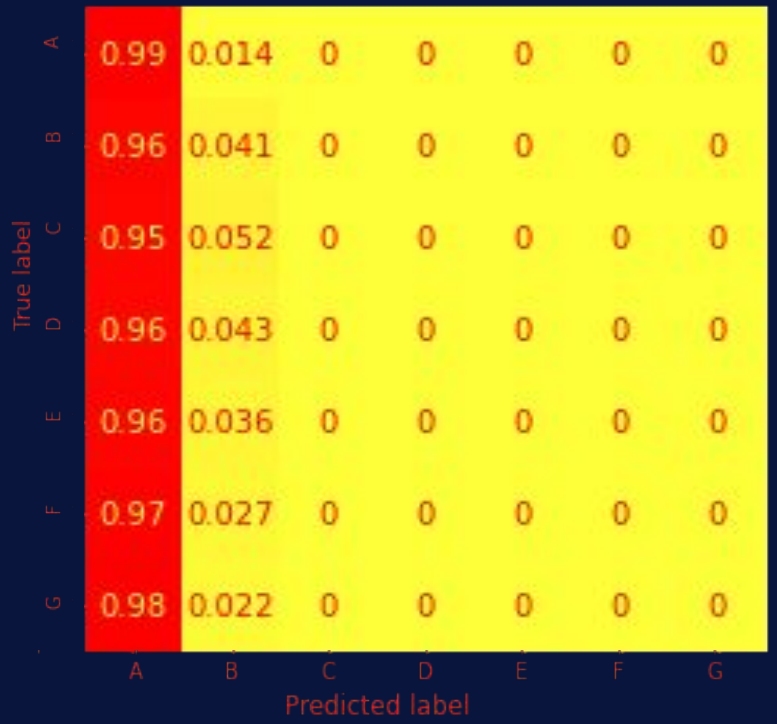
Most important feature (location excluded): Temperature average past 3 months

Second most important feature: Temperature average past 6 months

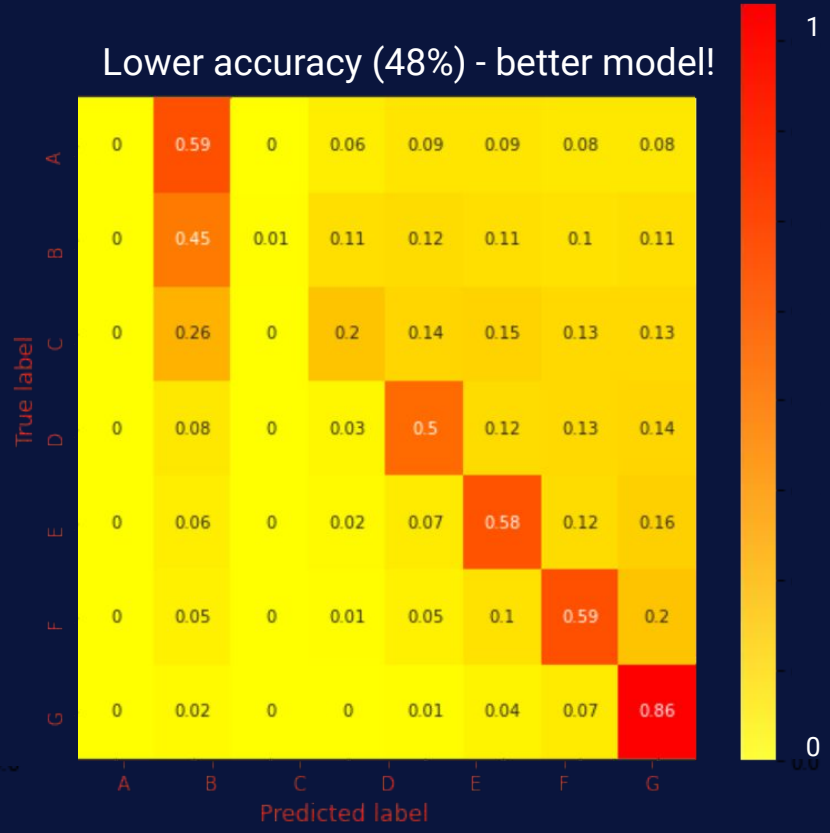
Third most important feature: Monthly Precipitation

WHAT DOES “GOOD” LOOK LIKE?

Higher accuracy (62%) - useless model



Lower accuracy (48%) - better model!



SUMMARY & NEXT STEPS

What we learned

- Location is important! ☑
- Temp & Precipitation have direct relationships ☑
- Bootstrapping helps with unbalanced classes ☑
- Focusing on recall helps us predict disastrous fires. ☑

What we would still like to explore

- Looking into Local Resources ☐
- Adding Altitude and Vegetation Data (Earth Engine) ☐
- Add Wind Data (NOAA) ☐

With Climate Change and Warmer Temps

More Risk

Dry and hot conditions will bring more fire risk.

Need for Security

We will need predictive models to protect communities.

More Resources

Being prepared ahead and being able to respond will be critical.

QUESTIONS?

THANKS!

DO FIRES GET LARGER AS CLIMATE GETS WARMER?