Scorched!

An analysis of
U.S. Wildfires
from 1992-2015

By

k.E.G.d

Katlin, Enzo,
Graham, and Danieli

Why Study Wildfires?

Per the Environmental Protection Agency,

- **Wildfires have the potential to harm property, livelihoods, and human health.**
- **Studies show that climate change has already led to an increase in wildfire season length, wildfire frequency, and burned area.**

Per the Colorado Division of Fire Prevention and Control - Top 5 all since 2018

Colorado's Largest Fires by Acreage

Rank	Fire	Acres	Year
1	Cameron Peak	208,913	2020
2	East Troublesome	193,812	2020
3	Pine Gulch	139,007	2020
4	Hayman	137,760	2002
5	Spring Creek	108,045	2018
6	High Park	87,284	2012

Images sourced from Google Images

Wildfires Close to Home

Marshall Fire, Dec 2021

RAW: Longmont Firefighter shares video of what it looked like fighting the Marshall Fire

Aerial video of Marshall Fire's destruction

East Troublesome Fire, Oct 2020

Images

Data Source

U.S. Wildfire data (plus other attributes)

Subset of 1.88 Million US wildfire joined with other related database

U.S. Wildfire data (plus other attributes)

https://www.kaggle.com/datasets/capcloudcoder/us-wildfire-data-plus-other-attributes

- Entire dataset - 55,367 rows / 43 columns
- Cleaned dataset - 13,138 rows / 18 columns

```
df
```

✓ 0.1s Python

	fire_id	fire_size	fire_cause	latitude	longitude	state	discovery_month	Temp_pre_30	Temp_pre_15	Temp_pre_7	Wind_pre_30	Wind_pre_15	Wind_pre_7	Hum_pre_3
0	3	1.00	Debris Burning	39.641400	-119.308300	NV	Jun	16.275967	18.996181	18.142564	4.054982	3.398329	3.671282	44.77842
1	24	40.00	Arson	31.435181	-88.999489	MS	Apr	13.468619	15.067227	15.604790	2.038268	1.737921	1.775904	57.99720
3	31	1.20	Debris Burning	48.833000	-99.783600	ND	Apr	-0.891635	0.372659	-4.273834	5.800667	6.012852	6.658621	77.57501
4	35	30.18	Debris Burning	31.259000	-84.895600	GA	Oct	20.079480	17.722714	18.188679	3.659840	3.366443	2.211429	67.55178
5	36	1420.00	Lightning	33.241800	-104.912200	NM	Jul	31.055859	32.523438	34.893333	4.026367	3.844922	3.695833	28.78320
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
15326	55336	3409.00	Utilities	31.059000	-98.956367	TX	Aug	30.731860	30.545367	28.719917	3.136761	3.238803	3.224274	50.08915
15327	55337	4582.00	Utilities	30.075167	-97.149167	TX	Oct	24.642268	23.713390	24.221869	1.529850	1.576828	1.563817	62.84817
15328	55341	17823.00	Accidental	44.834600	-117.220600	OR	Sep	15.546194	12.890633	10.734328	2.608150	2.486802	1.835821	55.00925
15329	55342	5086.00	Debris Burning	45.656100	-109.114200	MT	Mar	2.275974	7.360185	7.678571	4.428757	4.197593	3.803571	51.67668
15330	55343	5963.00	Debris Burning	43.785671	-115.985922	ID	Oct	17.007551	16.983670	15.732906	1.580406	1.835639	1.405983	51.54180

13138 rows × 18 columns

Images sourced fror

Questions to Investigate

- Has the number, size, and cause of wildfires changed over time? (general analysis and visualization)
- Has the timing and distribution of wildfires changed over time? (general analysis and visualization)
- Can temperature, wind, humidity, discovery month, and state be used to predict whether a fire will grow to at least 50 acres in size? (machine learning)

Images sourced from Google Images

Data PreProcessing: Newly Categorized Fire Size Bins

	count	fire_size_bin_no	fire_size_bin
0	5955	1	Teacup
1	1327	2	Toy
2	1982	3	Mini
3	1015	4	Medium
4	1610	5	Large
5	1249	6	XL

- Binned fire sizes to get a better view of proportions and more it easier to work within a machine learning models.
  - Teacup: 0-5 acres
  - Toy: 5-10 acres
  - Mini: 10-50 acres
  - Medium: 50-1,000 acres
  - Large: 1,000-10,000 acres
  - XL: 10,000 + acres

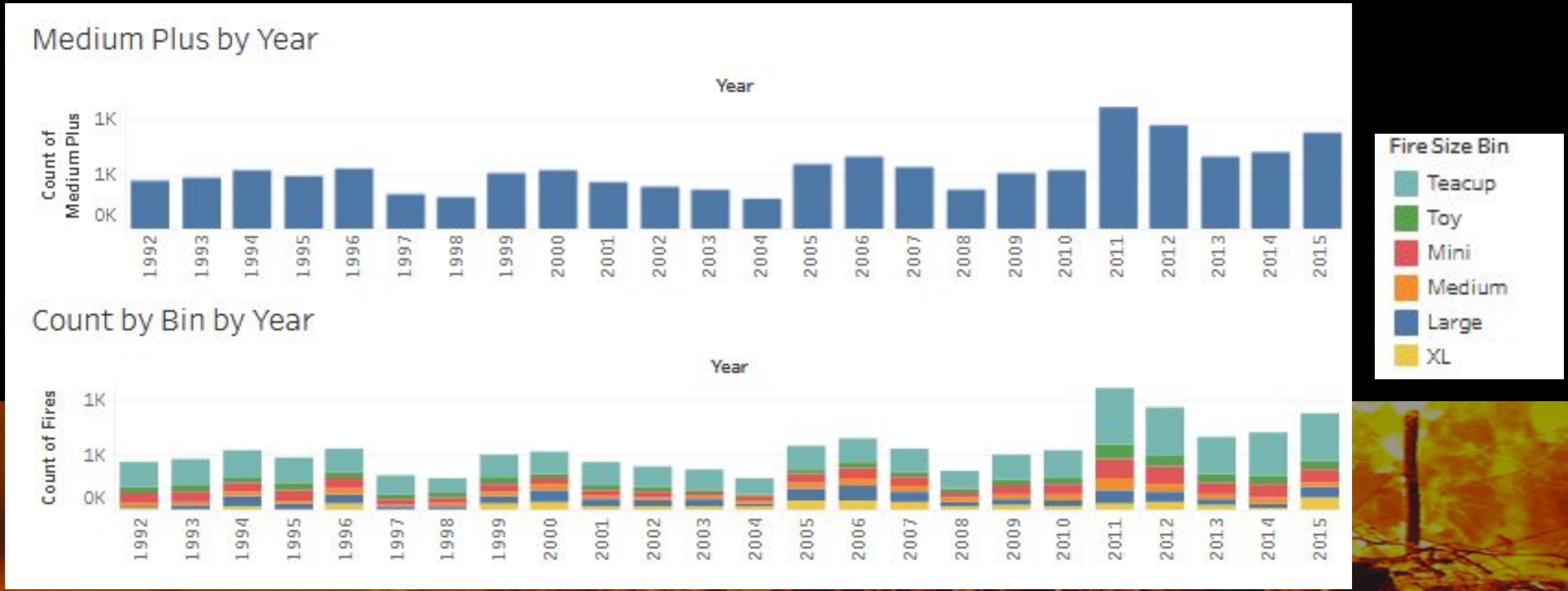
Images sourced from Google Images

Data Analysis: Bin Comparison by Year

	count	fire_size_bin	year
0	21	Large	1992
1	31	Medium	1992
2	88	Mini	1992
3	234	Teacup	1992
4	51	Toy	1992
5	20	XL	1992
6	91	Large	2015
7	47	Medium	2015
8	112	Mini	2015
9	419	Teacup	2015
10	85	Toy	2015
11	121	XL	2015

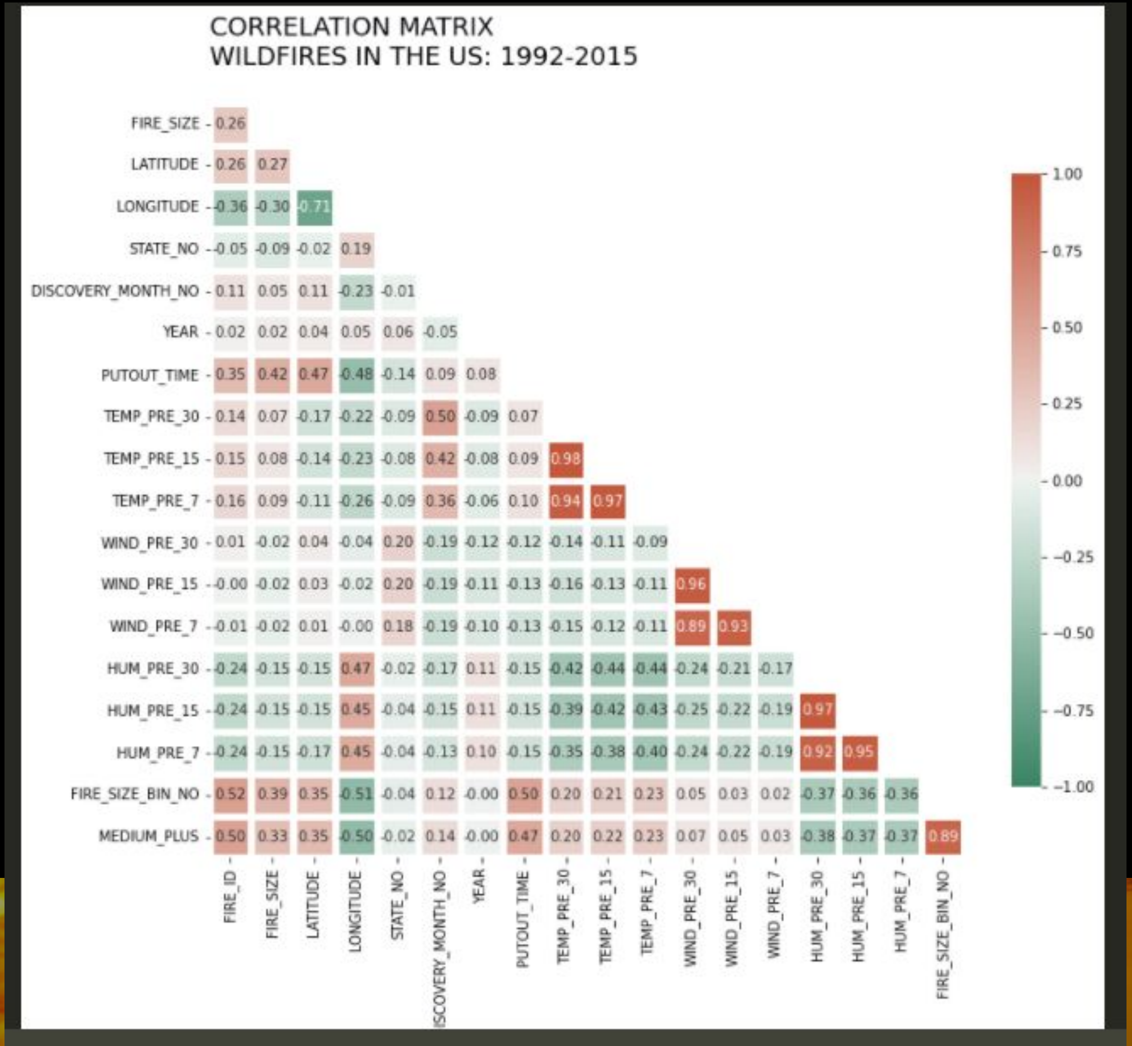
	year	count
0	2011	1103
1	2012	939
2	2015	875
3	2014	702
4	2013	661

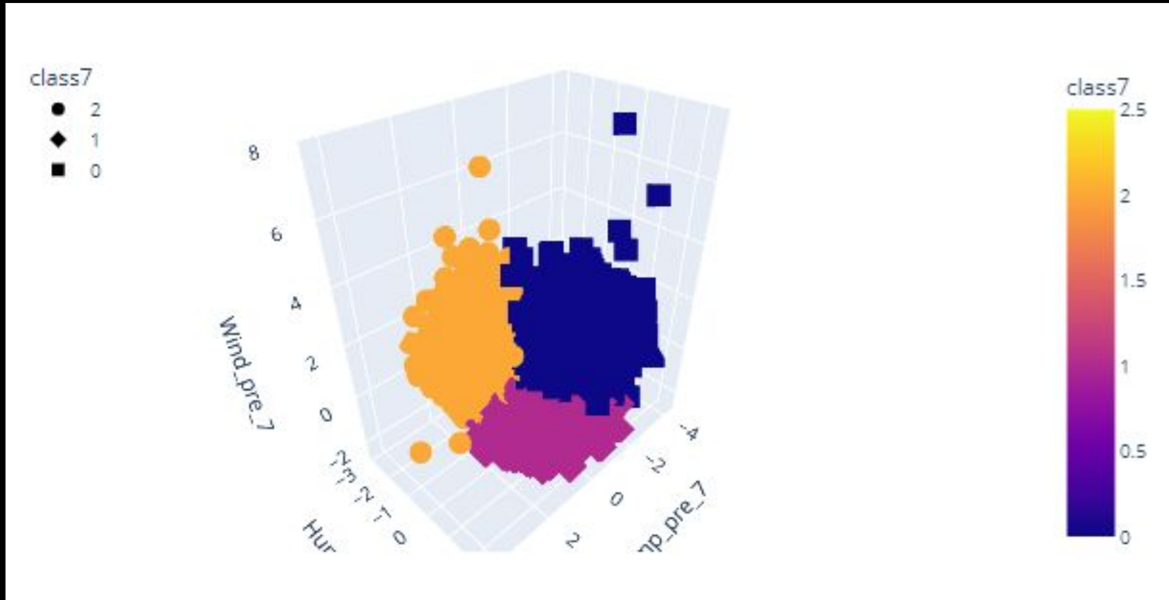
- Nearly all bin counts have increased between 1992-2015.
- Top 5 fire counts are last 5 years in study period.



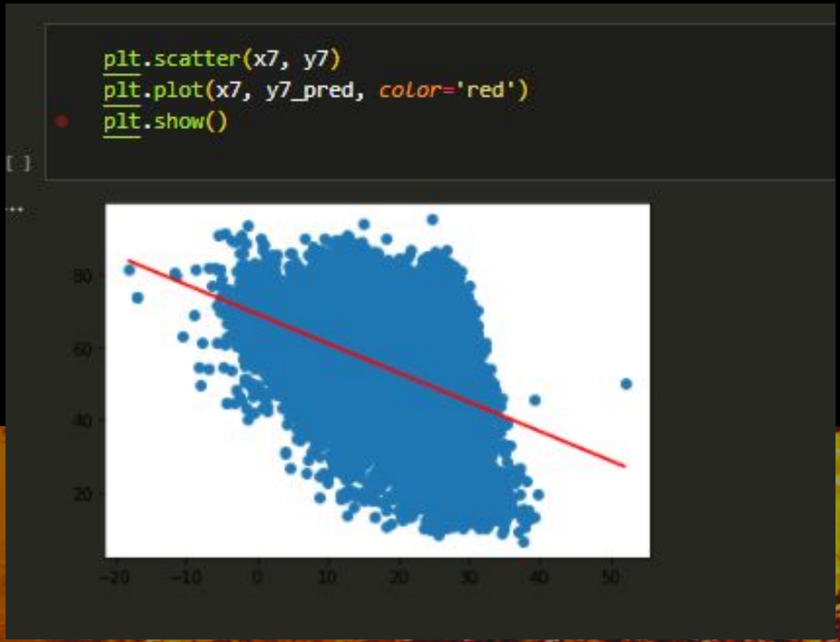
Images sourced from Google Images

Data Analysis: Other Charts





KMeans Clusters of Pre-7 Weather Data



Pre-7 Linear Regression of Temperature (x) vs Humidity (y)

Images sourced from Google Images

Data Analysis: <u>Supervised Machine Learning</u>

- **<u>Target (y) value</u>**: 'medium_plus'
- **<u>X values</u>**: 'state', 'discovery_month', 'Temp_pre_7', 'Hum_pre_7', 'Wind_pre_7'

One-Hot Encoder

```
Supervised Machine Learning Models Using OneHot Encoder
Neural Networking Accuracy: 0.7887367010116577
Random Forest Accuracy Score : 0.7237530346501876
Balanced Random Forest Accuracy Score : 0.7541105716177444
Easy Ensemble Accuracy Score : 0.7537486206135511
Naive Random Oversampling Accuracy Score : 0.7619487971750165
SMOTE Oversampling Accuracy Score : 0.7578845729419554
Centroid Clustering Undersampling Accuracy Score : 0.7499801368351358
SMOTEENN Over/Undersampling Accuracy Score : 0.7589009048775104
Gradient Boosting Accuracy Score : 0.7192584418450674
Logistic Regression Accuracy Score : 0.7796042617960426
```

Label Encoder

```
Supervised Machine Learning Models Using Label Encoder
Neural Networking Accuracy: 0.7001522183418274
Random Forest Accuracy Score : 0.717443169278305
Balanced Random Forest Accuracy Score : 0.7451909070845288
Easy Ensemble Accuracy Score : 0.7580284463894967
Naive Random Oversampling Accuracy Score : 0.7072702407002188
SMOTE Oversampling Accuracy Score : 0.7044901531728666
Centroid Clustering Undersampling Accuracy Score : 0.6951444201312911
SMOTEENN Over/Undersampling Accuracy Score : 0.7594628008752735
Gradient Boosting Accuracy Score : 0.7292002188183807
Logistic Regression Accuracy Score : 0.6138148311630987
```

Images sourced from Google Images

Analytical Conclusions

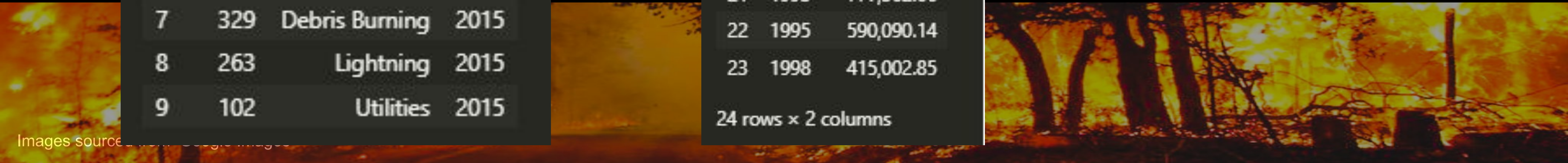
- The size of fires has steadily increased over the 23 year period.
- The total square acreage of burned area has hit a 23 year high in 2015.
- Increase in fire size may be a result of increase in temperatures during the study period.

	count	fire_cause	year
0	47	Accidental	1992
1	165	Arson	1992
2	116	Debris Burning	1992
3	96	Lightning	1992
4	21	Utilities	1992
5	73	Accidental	2015
6	108	Arson	2015
7	329	Debris Burning	2015
8	263	Lightning	2015
9	102	Utilities	2015

	year	sum
0	2015	5,956,913.82
1	2005	5,442,221.32
2	2004	4,797,085.48
3	2012	4,606,347.1
4	2006	4,556,531.91
...	...	...
19	1997	886,198.32
20	1992	814,850.27
21	1993	777,362.66
22	1995	590,090.14
23	1998	415,002.85

24 rows × 2 columns

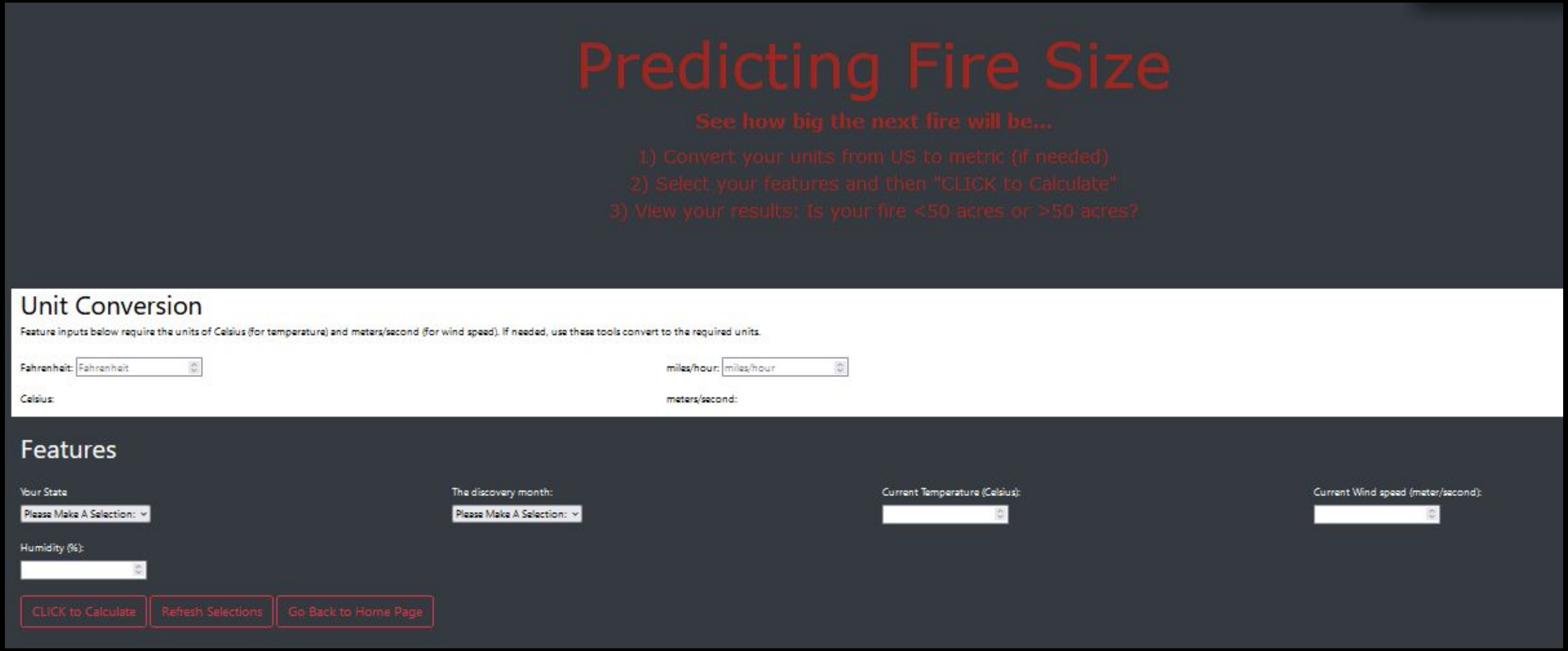
	max temp	year
0	34.361864	1992
1	38.146746	2015

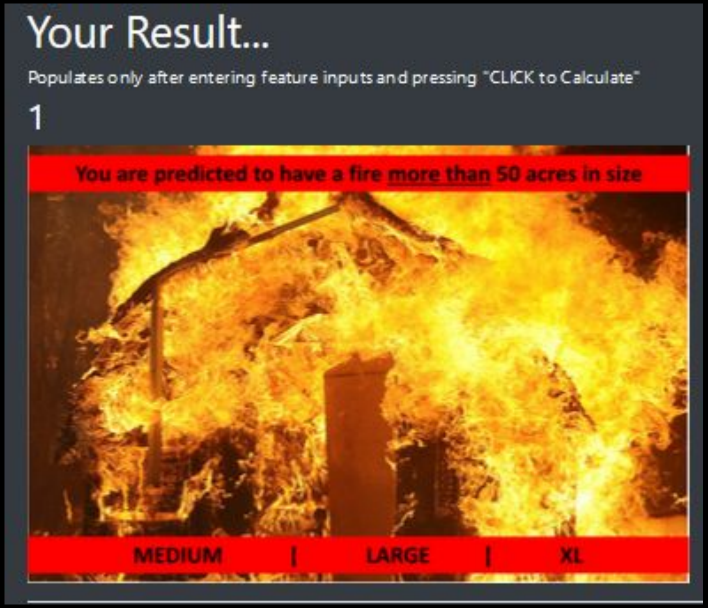


Images source

Predicting Fire Size: See How Big the Next Fire Will Be

Check out our Prediction Web App…



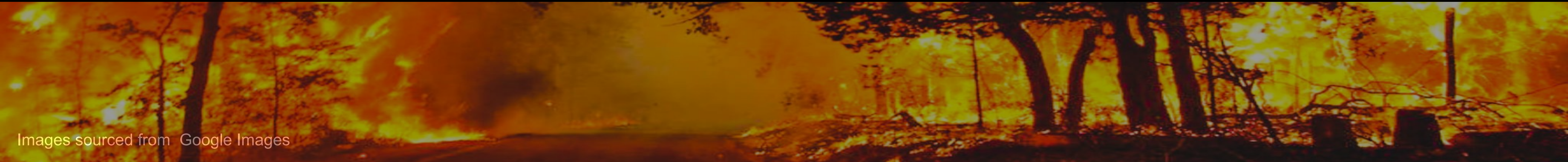


Images sourced from Google Images

Real Life Applications

With this fire prediction tool, first responders may address

- Asset relocation
- Early containment
- Possible evacuation areas
- **Prevent future loss of life and minimize property damage**
- Help with resource and funding requests
- Highlight highly affected areas that need attention



Images sourced from Google Images

Climate Change

As the climate change crisis now creates conditions that are deteriorating and causing worse destruction, having the capability to determine potential fire size is essential.

Climate change affects us all and needs to be a topic of discussion at all levels of government and public forum.

Tools like this one that are easy to understand and bring pertinent issues to the forefront are needed to bring us all in the fight for a more stable future.

Images sourced from Google Images

Credits / Thank Yous

Team k.E.G.d. wishes to thank our instructor, *__Svitlana__*, and our T.A.s, *__James, Sheri, and Simon__*, for all their assistance on this project.

Credit is also due to Team Lizard People, and their Conspiracy Theory project, for inspiration when our coding struggles got real.

Lastly, without DU's Data Analytics Bootcamp, this team, our friendship, and this project wouldn't even exist!!!

Images sourced from Google Images