Smoke Point

CALIFORNIA NEWSROOM TRAINING
JUNE 2022

BEFORE WE START …

- Your phones/computers are muted for the presentation
- The training will be recorded
- Unmute to ask questions at the end
- Questions also can be submitted via chat -- please send to “everyone”
- Please take a second to type your name, organization and email in the chat. We’ll invite you to our project slack channel. Thank you!

PRESENTERS

Mohar Chatterjee, Senior Research & Data Fellow, MuckRock, Brown Institute for Media Innovation (mc4958@columbia.edu)

EMBARGO, USE AND CREDITING

Our reporting team will be available for you in many ways.

They are also available to coach reporters who want to use the data to explore local or exclusive angles

Mandatory credit to “an analysis of federal satellite imagery by NPR’s California Newsroom and MuckRock and the Brown Institute for Media Innovation”

Media outlets outside the California Hub, please link back to the landing page for the story.

THREE DATA PRODUCTS

Census Tract Level Smoke Days + Demographic Data (national)

Regression Analysis (California + national)

Choropleth Maps (national)

Smoke in the Air

Smoke from raging wildfires in the West have grown in both intensity and frequency over the past decade, exacerbated by climate change and decades of fire-suppression policies.

Those most impacted by that wildfire smoke — and those suffering the health effects of frequent smoke exposure — have been hard to identify at the hyperlocal level.
Until now.

For the past six months, KQED, MuckRock, and Columbia University's Brown Institute for Media Innovation have been working together to understand how wildfire smoke exposure varies by neighborhood.

Quantifying Wildfire Smoke

- For the first time ever, we have created a hyperlocal dataset of the number of wildfire “smoke days” experienced by every census tract across the United States over the past 15 years.

- “Smoke days” refer to any day where a census tract resident would see a smoke plume produced by burning vegetation over their heads, as caught by NOAA satellites.

- These are best understood as frequency measures, with some additional information about intensity.

Some Topline Findings

We found that in many neighborhoods across the U.S., those primarily populated by residents most vulnerable to wildfire are also the ones who are most frequently exposed to it, even after controlling for a broad range of socio-economic factors.

E.g. We have identified 18 states, including California, Texas, Florida, New York, Michigan, Pennsylvania, where neighborhoods with a large proportion of residents in poverty are likely to be more frequently inundated with heavy wildfire smoke even after controlling for a broad range of other socio-economic factors.

Some Topline Findings

Across the state of California, the analysis also found that census tracts with a greater proportion of Native Americans were likely to experience far more wildfire smoke than neighborhoods largely populated by other races.

Some Topline Findings

In Northern California, the Bay Area, Central Valley and south to Ventura County, where the rise wildfire smoke exposure has been particularly severe, we found that neighborhoods with a greater proportion of poor and disabled residents were far more likely to be inundated with frequent wildfire smoke, even after controlling for a host of other socio-economic factors.

Look-up: Impact Analysis by State

How to Interpret Coefficients: Exponentiate the coefficient, subtract one from this number, and multiply by 100. This gives the percent increase (or decrease) in the response for every one-unit increase in the independent variable. Example: the

State_Abbr	Regression	Household	HBPL_Coef	People_w	PWD_Coef	People_65	P65_Coeff	People_w	PWESD_Co	Household	HWNV_Co	Hispanic_o	HOL_Coeff	White_Per	W_Coeff	Black_or_A	BOAA_Coe	American_	AIAN_Coef	Asian_Perc	A_Coeff
WA	YES	YES	0.004172	NO		NO		YES	0.004124	NO		NO		NO		NO		NO		NO	
DE	YES	NO		NO		NO		NO		NO		NO		NO		NO		NO		NO	
WI	YES	YES	0.002199	NO		YES	0.001438	YES	0.005657	NO		NO		YES	0.00987	NO		YES	0.011001	NO	
WV	YES	NO		NO		NO		NO		YES	0.003321	NO		NO		NO		NO		YES	0.020339
HI	YES	NO		NO		YES	0.042055	NO		NO		NO		NO		NO		NO		NO	
FL	YES	YES	0.009098	YES	0.033503	NO		NO		NO		NO		YES	0.031749	YES	0.02859	YES	0.083191	NO	
WY	YES	NO		NO		YES	0.008657	NO		NO		NO		NO		NO		NO		NO	
NJ	NO																				
NM	YES	NO		YES	0.015513	YES	0.002803	NO		NO		NO		NO		NO		NO		NO	
TX	YES	YES	0.003007	YES	0.008423	NO		YES	0.008496	NO		NO		YES	0.020117	NO		YES	0.097473	NO	
LA	YES	YES	0.007996	NO		YES	0.005617	NO		NO		NO		NO		NO		NO		NO	
NC	YES	YES	0.001996	YES	0.008732	NO		NO		NO		NO		NO		NO		NO		NO	
ND	YES	NO		NO		YES	0.002032	NO		NO		YES	0.018088	YES	0.020078	YES	0.018657	YES	0.02052	YES	0.019763
NE	YES	YES	0.004243	NO		YES	0.006852	NO		NO		NO		NO		NO		NO		NO	
TN	YES	NO		YES	0.013957	NO		NO		NO		NO		NO		NO		NO		NO	
NY	YES	YES	0.003045	YES	0.00153	NO		NO		NO		NO		NO		NO		YES	0.002729	NO	
PA	YES	YES	0.001329	YES	0.005715	NO		YES	0.00176	YES	0.00126	NO		NO		NO		NO		NO	
AK	YES	NO		NO		NO		NO		YES	0.009666	NO		NO		NO		NO		NO	
NV	YES	NO		YES	0.017742	NO		NO		NO		YES	0.03195	YES	0.041475	NO		YES	0.05714	NO	
NH	YES	NO		NO		NO		YES	0.003848	NO		NO		NO		NO		NO		NO	
VA	YES	NO		NO		NO		NO		YES	0.001246	NO		NO		NO		NO		NO	
CO	YES	NO		NO		NO		YES	0.003589	NO		NO		NO		NO		NO		NO	
CA	YES	YES	0.011517	YES	0.010155	NO		NO		NO		NO		NO		NO		YES	0.030118	NO	
AL	YES	NO		YES	0.01383	NO		NO		NO		YES	0.032883	YES	0.036189	YES	0.033803	YES	0.036883	NO	
AR	YES	NO		NO		NO		NO		NO		NO		NO		NO		YES	0.031853	NO	
VT	YES	NO		NO		NO		NO		NO		NO		NO		NO		NO		NO	
IL	YES	NO		NO		YES	0.00073	NO		NO		NO		NO		NO		NO		NO	
GA	NO																				
IN	YES	NO		NO		YES	0.003327	NO		YES	0.001858	NO		NO		NO		NO		NO	
IA	YES	NO		NO		YES	0.002331	NO		NO		YES	0.015281	YES	0.014209	YES	0.013798	YES	0.016777	NO	

The Smoke Plumes Dataset

- Total light, medium, and heavy density smoke days experienced by each census tract in the United States per year from 2006 to 2021. There are over 84,000 census tracts in the United States.

- Geometries for smoke plumes come from NOAA satellite imagery, specifically the GOES Aerosol and Smoke Product (GASP). Smoke density for each smoke plume is measured by looking at the aerosol depth.

- To ensure we only counted days with noticeable smoke exposure for neighborhood residents, we filtered our dataset to only look at medium or heavy density smoke days.

The Smoke Plumes Dataset

–Smoke Metrics:

(1) the average number of annual smoke days experienced by each census-tract over the past 6 years.

(2) The interpolated rate of increase in the number of annual smoke days experienced by each census-tract over the past 6 years.

We chose the 6 year time period to ensure our data was not too skewed by years with too many or too few wildfires, while also remaining relevant to our reader's recent experiences.

The Analysis

–Based on CalFire's criteria for wildfire vulnerability, we collected ACS data on the following for each census tract (demographics reported as normalized percentages):

- People living with disabilities
- Households without a vehicle
- Households whose income is below the poverty line
- People who are 65 years old and above in age
- People who have difficulty speaking or understanding English
- People of each race living in each census tract.

An Example

CALIFORNIA

OLS Regression Results

Dep. Variable:	natural_log_ROC	R-squared:	0.113
Model:	OLS	Adj. R-squared:	0.112
Method:	Least Squares	F-statistic:	84.79
Date:	Fri, 06 May 2022	Prob (F-statistic):	1.91e-197
Time:	01:41:13	Log-Likelihood:	-1035.9

	coef	std err	P>|t|
const	1.1498	0.492	0.020
Households_below_Poverty_Level_Proportion	1.0332	0.047	0.000
People_w_Disabilities_Proportion	1.3126	0.202	0.000
People_65_n_Over_Proportion	-0.1094	0.050	0.028
People_w_English_Difficulty_Proportion	-0.2451	0.048	0.000
Households_w_No_Vehicle_Proportion	-0.0966	0.044	0.027
Hispanic_or_Latino_Proportion	0.3682	0.493	0.456
White_Proportion	0.6079	0.494	0.218
Black_or_African_American_Proportion	0.4180	0.498	0.401
American_Indian_Alaska_Native_Proportion	1.7610	0.529	0.001
Asian_Proportion	0.5232	0.494	0.290
Native_Hawaiian_Pacific_Islander_Proportion	1.6433	0.590	0.005
Two_or_More_Races_Proportion	1.5913	0.523	0.002

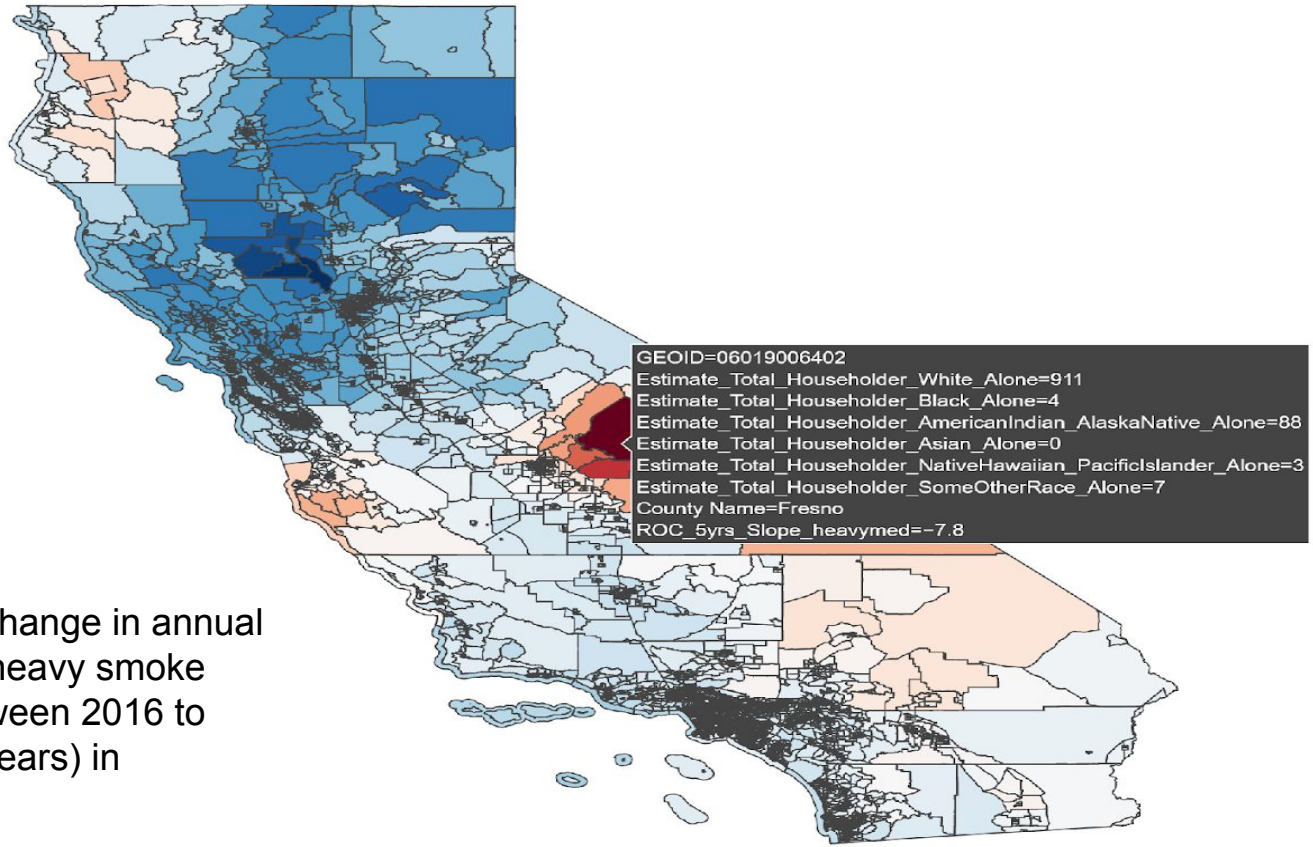
Behind the Scenes

- Based on advice from our expert panel, we log-transformed our two smoke metrics for 6 year average and 6 year rate of change — these became our dependent Y variables. For each Y variable in turn, we ran a state-wide regression using these census-tract level data-points.

- We found some positive, statistically significant correlations between frequency of smoke exposure to neighborhood-level socioeconomic demographics that indicate wildfire vulnerability.

- Our minimum threshold for statistical significance for any regression model was an adjusted R^2 value of at least 10%, and for a coefficient is only considered significant if its associated p-value is less than 5%.

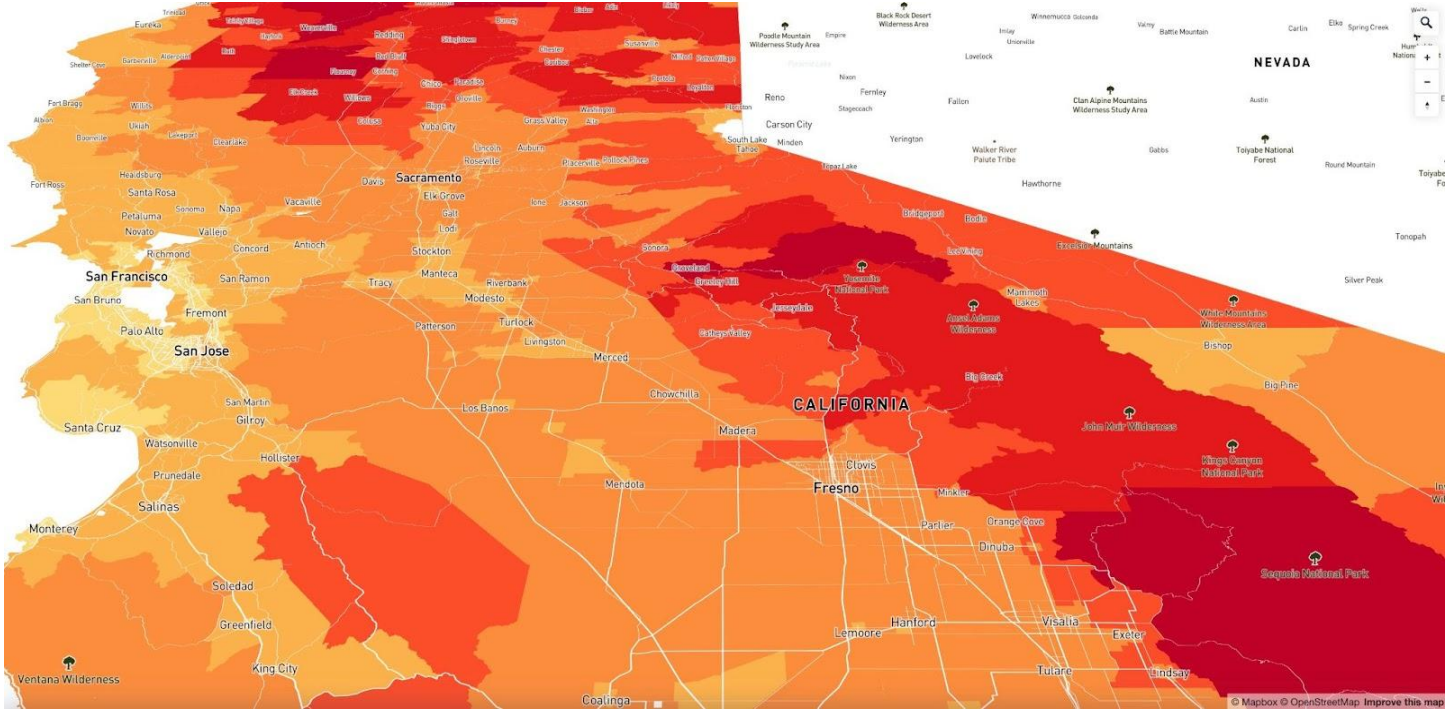
Rate of Change Smoke Days Choropleth Map



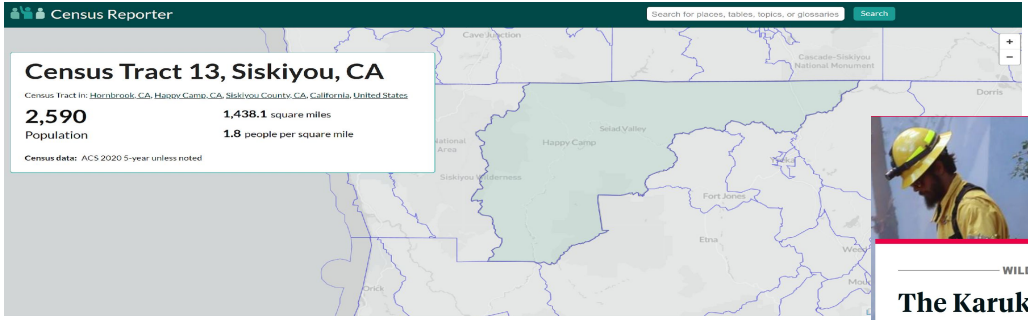
Rate of Change in annual medium-heavy smoke days between 2016 to 2020 (5 years) in California

Average Smoke Days Choropleth

Mean count of annual medium-heavy smoke days between 2016 to 2021 (6 years) in California



A Quick Walk-Through



WILDFIRE

The Karuk Used Fire to Manage the Forest for Centuries. Now They Want To Do That Again

LISTEN

By Danielle Venton May 28, 2021 Save Article

For thousands of years before contact with Europeans, the Karuk people, like many Native American tribes, tended their land with fire, keeping an ecological balance among plants, animals, river, and forest.

Situated along the Klamath River, in Humboldt and Siskiyou counties, the Karuk are now struggling to renew their way of life.

Today, nearly 98% of the tribe's ancestral land is controlled by the U.S. Forest Service. The landscape is overgrown with timber and undermanaged. Out-of-control wildfires have repeatedly decimated the area, a consequence of disallowing native people to wield the tool of intentional burning, say advocates of the practice.

California

Fire tore through the Karuk tribe's homeland. Many won't be able to rebuild

Vivian Ho *in Happy Camp, California*

Fri 23 Oct 2020 06.00 EDT

Erin Hillman, a member of the Karuk Tribe, looks at ceramics damaged in her home after the Slater fire in Happy Camp, California, on 30 September 2020. Photograph: Carlos Avila Gonzalez/The San Francisco Chronicle via Getty Images

What’s going on (Speaker Notes)

For example, take the choropleth of mean smoke days in California. That represents every census tract in California, color coded by the mean smoke days they've seen in the past 6 years, and embedded with tons of information on each neighborhood upon hover. A reporter could use that to zoom in a remote census tract in the northernmost part of the state, say Siskiyou County, that's consistently seen a lot of smoke. Turns out that tract has a rather high number of Native American residents. A 5 minute Googling session will tell you that tract contains the headquarters of one of the largest tribes in California - the Karuk Tribe. Going to the regression for California, I'd also see a statistically significant correlation between neighborhoods with high smoke exposure and neighborhoods with Native American residents, meaning that tract was not an anomaly - but rather, part of a larger story. That story took me a total of 20 minutes to find and confirm. Shoe leather reporting is needed to flesh out the story, but you'll notice that our key contribution to the newsroom here is reducing their lead generation time. That local reporter has a neighborhood to focus on, and a story framework they can work off of. That's pretty valuable to our newsroom partners. And for the audience, we're able to explain on a hyperlocal level, the connection between climate related phenomena to the socio-economic dynamics of their neighborhood - connections they may already see and perceive, but now they will be able to contextualize their experience.

Finding your Area and Relevant Analysis:

- Link to Spreadsheets.
- Links to Smoke Days Code:
  - For all states: https://colab.research.google.com/drive/1M_gpo3c1qNXLlzy4LteDxH9ey3SB6vUk?usp=sharing
  - For California: https://colab.research.google.com/drive/1GwAy86TXUGatL_CE17psxbZpmZ9oOwXx?usp=sharing
  - For Regressions: https://colab.research.google.com/drive/1MUvioY_5zisWPp8Uj7PaPswflDsL9Rop?usp=sharing

Reach out to:

- Researchers studying public policy, climate change, and environmental health:
  Dr. Francesca Dominici, fdominic@hsph.harvard.edu
  Dr. Colleen Reid, colleen.reid@colorado.edu
  Dr. Rosana Aguilera, r1aguilerabecker@ucsd.edu
  Dhrumil Mehta, dm3494@columbia.edu

- Pretty much anyone! School children, parents, farmers, ranchers, doctors, athletes, the elderly

- We are available

CAVEAT

- Regression Analysis results must be communicated carefully in order to preserve statistical integrity of our findings.
- Wording in previous slides has been vetted for accuracy and can be used as examples.
- **E.g.:** *“For every additional percentage of [INSERT SOCIO-ECONOMIC FACTOR] in a neighborhood, the residents will experience an additional [INSERT COEFFICIENT * FORMULA] of medium to heavy wildfire smoke on average, after accounting for a host of other race and socio-economic factors.”*
- Feel free to check in with us if you have questions about your region

CONTACTS

Mohar Chatterjee
mc4958@columbia.edu